

PS/G2 - 2

THE BRITISH JOURNAL OF STATISTICAL PSYCHOLOGY

EDITED BY
JOHN WHITFIELD

WITH THE ASSISTANCE OF
CYRIL BURT

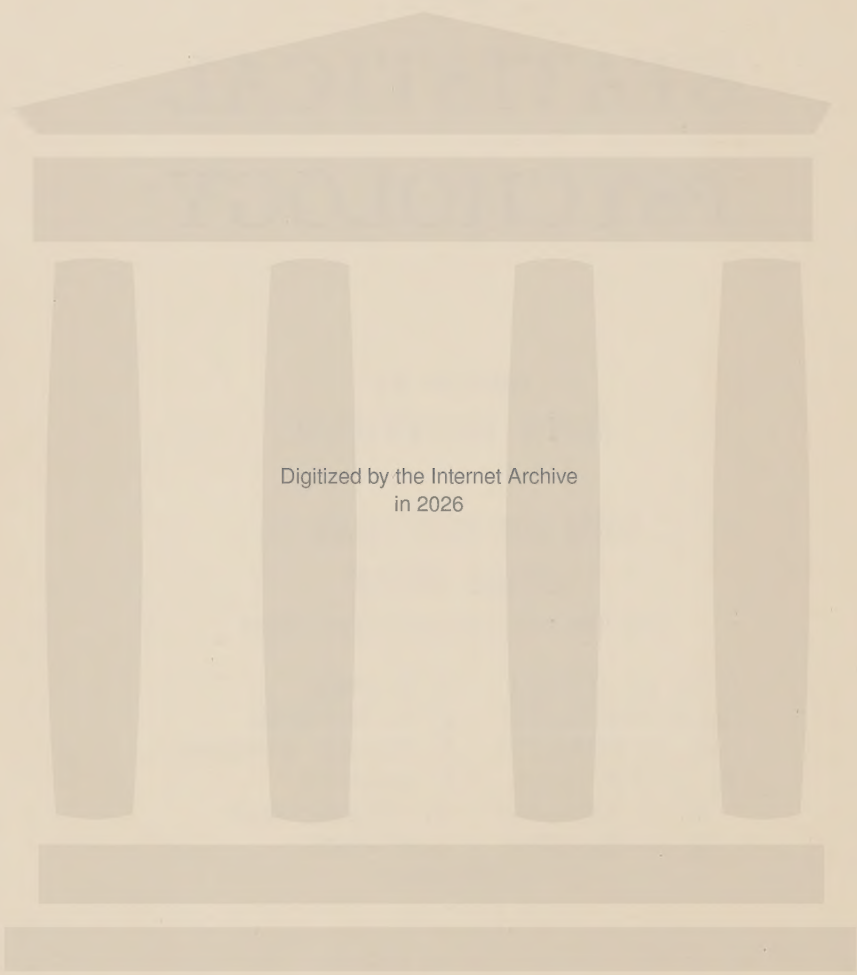
AND THE FOLLOWING EDITORIAL BOARD

A. C. AITKEN	E. A. PEEL
M. S. BARTLETT	L. S. PENROSE
W. G. EMMETT	J. FRASER ROBERTS
M. G. KENDALL	A. RODGER
D. N. LAWLEY	W. STEPHENSON
E. S. PEARSON	A. STUART
P. E. VERNON	

VOLUME XIII

Printed and Published by
TAYLOR & FRANCIS LTD.

RED LION COURT, FLEET STREET, LONDON, E.C.4



Digitized by the Internet Archive
in 2026

GUSTAV THEODOR FECHNER
ELEMENTE DER PSYCHOPHYSIK

1860

"Let us now praise famous men" . . .
For their work continueth,
Broad and deep continueth,
Greater than their knowing.

Rudyard Kipling, *Stalky & Co.*

The centenary of the publication of Fechner's *Elements of Psychophysics* is an occasion which the statistical psychologist can hardly overlook. The student of psychology commonly thinks of Fechner as "the first man who persuaded people to keep on lifting weights". William James tells us that "his book was the starting point of a new department of literature, of which, in the humble opinion of the present writer, the psychological outcome is just *nothing*". Flügel, on the other hand, in his *Hundred Years of Psychology* declares that "the *Elements* marked the birth of a specifically quantitative psychology, which has sought, with increasing if not with perfect success, to give to the mental sciences something of the same exactitude as the physical sciences, by adopting their methods". In embryo the idea had already been maturing in the writings of Herbart, Helmholtz, Müller, and Weber; but, "if Weber was, as Fechner styles him, 'the father of psychophysics', Fechner was the midwife, who successfully brought the lusty infant to birth".

James' verdict on the famous psychophysical law as a mere "idol of the den if ever there was one" and his picture of the "dear old man with his patient whimsies" still seem to dominate the current impressions of Fechner and his work. James, however, had little sympathy with quantitative analysis or statistical inquiries. And to get a truer conception of Fechner's service to psychology it is necessary, I suggest, to recall something of the background of his various researches.

Fechner was by no means the man of a single book and a single idea as is commonly assumed. Wundt's edition of the *Elemente* contains a bibliography of 175 publications issued under Fechner's name or that of 'Dr. Mises'. Fechner indeed had all the versatility of the typical German *Gelehrte*. He was by turns a physiologist, a physicist, a philosopher with a mission, a neurotic invalid, a psycho-physicist, an experimental aesthetist, and, in his leisure moments, a satirist and poet.

Fechner was born in 1801 in a small village near the borders of Saxony and Silesia in South-East Germany. His father was the village pastor, noted

for his own original ideas : he scandalized his flock by erecting a lightning-conductor on the steeple of his church, "as if God was unable to protect his own house", and by preaching without a wig, which, so he announced, "Jesus himself never wore". Gustav was educated at a nearby *Gymnasium* ; and at sixteen matriculated as a student of medicine in the University of Leipzig, where Weber was already professor of Anatomy. And at Leipzig he passed the remainder of his long life.

In those days Germany still looked to France for the latest work in science ; and Fechner, after graduating, supported himself by translating Thénard's six-volume textbook on chemistry, Biot's five volumes on physics, and Léon Rostans' work on diseases of the brain. In 1824 he was appointed lecturer in physics ; and a few years later, after a long series of experiments suggested by Ohm's recent 'law', he published a work on quantitative measurements of galvanic currents, which at once made his reputation as a physicist.

In 1834, at the early age of 33, he was invited to accept the chair of physics, and thus became full professor in the university where Herbart had preceded him and Wundt followed some thirty-six years later. During the next five years, in addition to experimental researches on chemistry and 'the theory of galvanism' (largely inspired by Faraday's new discoveries), he published several highly original psychological studies—including some still cited as landmarks, such as those on after-sensations, complementary colours, subjective colours, and the effects of 'primary memory'. At length, a temporary blindness caused by experiments with bright lights, and a prolonged nervous breakdown due to overwork, compelled him to relinquish his professorial duties for the next few years.

During this period of enforced relaxation his mind became more and more preoccupied with philosophical problems. Under the pseudonym of 'Dr. Mises' he had already begun, while still a youth of twenty, to publish occasional satires on the materialistic tendencies evinced by the physiological writers of his day. An ironical 'Panegyric of Contemporary Medicine' was followed by a witty burlesque, entitled *Vergleichende Anatomie der Engel*, and the more serious *Das Büchlein vom Leben nach dem Tode* (which ran through three editions in his lifetime), and later on by *Nanna oder über das Seelenleben der Pflanzen*. As these earlier contributions show, Fechner was greatly attracted by the romantic *Naturphilosophie*, which came trailing clouds of speculative fancy across the valley of dry bones left by Kant's *Critiques*. In particular Schelling's *Identitätslehre* (identity of subject and object) and his exploitation of the principle of 'bipolarity' (borrowed from the new theories of electricity and magnetism) seemed to him to be groping in the right direction. And yet his own training in physics and biology rebelled against the complacent disregard of scientific precision and empirical evidence. "Could any of the German *Naturphilosophen*," he exclaims, "ever have evolved anything so orderly and so convincing as the physical and chemical theories of contemporary French savants?"

For ten years he published comparatively little¹. But on the morning of the 22nd October, 1850, he tells us, while lying in bed, he had "a vision of a unified world of spirit and matter, linked by the mystery of numbers". Early in the following year he published what he regarded as his most important philosophical work, *Zend-Avesta, oder Über die Dinge des Himmels und des Jenseits: vom Standpunkt der Naturbetrachtung*.

The West had recently started rediscovering the East. Excerpts from the *Gathas*, the *Upanishads*, and the Buddhist scriptures had been freely translated into German. Schelling, Schlegel, and Hegel are full of allusions to oriental speculations. Schopenhauer christened his poodle 'Atma' ('Soul of the world'), and kept a bronze statuette of Buddha facing a plaster bust of Kant on his desk. Rückert, professor of oriental languages at Berlin, was publishing his 'Nachahmungen und Übersetzungen' of Indian and Persian poems². Wagner was meditating an Indian theme for a forthcoming opera, with cosmic implications, which eventually got embodied in the love-duet and the Verklärungs-aria of *Tristan und Isolde*.

To Fechner Indian pantheism supplemented by a Persian dualism—both, so he believed, synthesized and reconciled in the teaching of Zarathustra—seemed to offer a cosmology far more satisfying than the theology of his early childhood or the materialism of Laplace and Cabanis or Büchner and Vogt. His guiding idea was to substitute for the thoroughgoing idealism of the Hegelian philosophy and the out-and-out materialism of its scientific opponents a doctrine of psychophysical parallelism which, he held, could be applied both to the 'cosmos' and to the 'microcosmos', i.e., to man. Schopenhauer had flippantly declared that "ein denkendes Wesen ohne Gehirn ist wie ein verdauendes Wesen ohne Magen". Fechner substitutes 'Geist' for 'Gehirn', and adds that each of the two statements embodies a partial aspect of the truth. In keeping with his demand for supporting even philosophical speculations by scientific facts, we find in the middle of the work an announcement that he has discovered a simple mathematical relation between the spiritual and the material world, and a systematic programme of the psychophysical experiments which are to confirm it.

The full theory was set forth, with detailed evidence, in the *Elemente*, published some ten years later, when Fechner was already an elderly man of

¹ German philosophy at this juncture was in some measure a reflection of its political problems. In 1848 the popular demand for closer economic and political unification of the thirty-nine German states, and the agitation for a more liberal constitution, which followed the revolutionary outbreaks in France and elsewhere, stirred Fechner for a while into renewed activity. The outcome was several papers on topics not usually associated with his name: 'Über Volkssouverainetät', 'Über das Lustprincip des Handelns' and 'Über directe und indirecte Wahlen der Volks-Abgeordneten'. These were followed by a short article on the insistent problem of organization in a very different field—'Die mathematische Behandlung organischer Gestalten'—all published within sixteen months.

² 'Dr. Mises' had already printed critical notices of Rückert's poems in 1835. Rückert himself is now remembered, not for his oriental studies, but for his lyrics and songs, immortalized by Schubert, Schumann, and Brahms.

sixty. The aim of 'psychophysics', we are told, is to produce "an exact theory of the functional- or dependency-relations (funktionellen oder Abhängigkeitsbeziehungen) between body and mind and, more generally, between the physical and the psychic world". But the two are to be regarded, not as *merely* parallel activities—"like those of Leibniz's two clocks, ticking side by side on the same shelf", but rather as the inner and outer aspects of a single sphere. "Das was in der Sache Eins ist, von zweierlei erscheint." Had he perhaps in mind Rückert's well-known translation of the Brahminal proverb?

Zwei Hälften machen zwar ein Ganzes, aber merk :

Aus halb und halb getan entsteht kein ganzes Werk¹.

Fechner's main argument was not unlike that developed about the same time by Bain. There was already increasing evidence that physical energy is strictly conserved ; hence there could be no interaction or exchange of energy, but only, in Bain's phrase, "a correlation of physical and mental forces"². The key to the relation, so Fechner's intuition led him to believe, could be found by "treating the *relative* increase in bodily energy as the equivalent of the *absolute* increase of the corresponding mental intensity", in other words, a geometric series of physical energies would be paralleled by an arithmetic series of mental intensities. This implied that "the mathematical relation is logarithmic"³.

The essential part of the *Elemente* is concerned with what Fechner calls 'External Psychophysics', i.e., the relations between stimulus and sensation. The measurement of the physical stimuli presented no difficulty. His first real problem therefore was to devise a means for measuring the corresponding sensations. After an introductory section expounding his ultimate aim, he begins by explaining the concept of 'thresholds'—both 'absolute' and 'differential', and then describes the three celebrated 'psycho-physical methods' in detail, with the mathematical formulae appropriate to each, and the requisite tables of the normal curve of error. The longest section of the book is concerned with a systematic presentation of the 'fundamental facts'. With the aid of his brother-in-law, Professor A. W. Volkmann, and his own two nephews, he had spent five arduous years carrying out his elaborate experi-

¹ 'Sprüche', *Aus der Weisheit der Brahmanen* (1828).

² Bain's arguments are set forth in detail in his *Mind and Body* (1872) and in his later appendix to Balfour Stewart's *Conservation of Energy* (1874).

³ Fechner seems only to have become familiar with the results of E. H. Weber's experiments (*Tastsinn und Gemeingefühl*, 1846) after hitting on his own logarithmic formulation. Weber, however, had merely shown by rather rough and ready methods that, for the senses studied, the just noticeable difference in sensation bore a constant ratio to the stimulus. He regarded it as a physiological not a psychological or psychophysical phenomenon ; and, as Gardner Murphy observes, the data obtained warranted no "sweeping generalization". Weber's greatest contribution lay in turning the physiologist's interest from the higher senses of sight and hearing to the lowlier senses of touch, temperature, and kinaesthesia. Thus, what Fechner generously named 'Weber's law', was really his own. Weber's empirical generalization of course has the merit of avoiding the highly controversial assumptions involved in Fechner's final formula.

mental programme : in the investigation of weight-discrimination, for instance, over 67,000 comparisons were made. Original data are given and discussed (i) for 'intensive thresholds', in the case of light and colour, loudness and pitch, weight and temperature, and (ii) for 'extensive thresholds' in the case of sight, touch, time, and movement ; and the relevant literature of physics, astronomy, and physiology is ransacked for scattered results supporting or conflicting with his own conclusions.

In the second volume 'the general mathematical function' ($S = K \log R/R_0$) is deduced and explained at length. The work concludes with an exposition of 'Internal Psychophysics', i.e., of the relation of sensation, not indirectly to the stimulus, but directly to the underlying 'excitation'. Here Fechner argues that his formula, which is only approximately satisfied in 'External Psychophysics', should be fulfilled exactly. He notes that its form implies the existence of 'negative sensations', corresponding to magnitudes of the stimulus below the threshold-value ; and he held that this offered an argument for a theory of unconscious mental states, not unlike that already adumbrated by Leibnitz and Herbart. Various practical applications of his law are noted : e.g., its implications for the classification of stars according to magnitude¹ and its reappearance at higher mental levels, e.g., in the field of motivation and of economics². He therefore suggests that his law may prove as fundamental in

¹ Hipparchus (c. 150 B.C.) in his catalogue had classified all the visible stars into six classes according to their apparent brightness. Early in the nineteenth century John Herschel and others carried out photometric measurements of the brightness of typical stars in each class ; and Herschel himself concluded that, while the apparent brightness of the several classes diminished by equal amounts, the photometric measurements were proportional to the reciprocals of the squares of integral numbers, i.e., to $1/1, 1/4, 1/9, 1/16$, etc. Fechner maintained that, with a more accurate choice of typical stars, the photometric values diminished in a geometrical series, i.e., that the objective measurements were proportional to $1/2, 1/4, 1/8, 1/16 \dots$ or (more precisely) to $1/2\frac{1}{2}, (1/2\frac{1}{2})^2, (1/2\frac{1}{2})^3$, etc. The modern grading, which requires fractional magnitudes for stars of intermediate brightness (as indeed Fechner suggested) is based upon Fechner's own principle. A star of the 1st magnitude is defined as one whose brightness is exactly 100 times that of a star of the 6th magnitude. As there are 5 intervals, an interval of one magnitude is equivalent to $\sqrt[5]{100}$; and, since $1/5 \log 100 = 0.4$, and $0.4 = \log 2.512$, a star of the 1st magnitude, like Aldebaran (the 'eye' of the Bull) should be approximately $2\frac{1}{2}$ times as bright as a star of the 2nd magnitude, such as the Pole Star. With a natural extension of the scale Vega, which is $2\frac{1}{2}$ times as bright as Aldebaran, is consequently assigned a magnitude of zero (or more exactly 0.14) ; Sirius is given a calculated magnitude of -1.6 , while those of Venus, the full moon and sun work out at $-4.0, -12.5$, and -26.7 respectively.

² It may be noted that, in the field of motivation, both Daniel Bernoulli and Laplace had already hit on the Weber-Fechner formulae. "Valde probabile est" says Bernoulli, "lucrum quodvis semper emolumentum afferre summae bonorum reciprocae proportionale", and he accordingly derives both the differential and the logarithmic formulae (*Acad. Scient. Imp.*, Petersburg, 1738, pp. 177-182). Laplace similarly distinguishes 'relative (or motivational) value' from 'absolute value': "D'après ce principe", he writes, " x étant la fortune physique d'un individu, l'accroissement, dx , qu'elle reçoit, produit à l'individu un bien moral réciproque à cette fortune; l'accroissement de sa fortune morale peut donc être exprimée par $k dx/x$, k étant une constante. Ainsi en désignant par y la fortune morale correspondante à la fortune physique x , on aura

$$y = k \log x + \log h,$$

h étant une constante arbitraire" (*Théorie analytique des probabilités*, 1812, p. 432).

psychology as the law of gravitation in general physics. "Doch sind dies für jetzt nur Ansichten und Aussichten."

The work remains one of the classics of psychology, and, like most classics, is now chiefly familiar from second- or third-hand accounts. Almost from the outset it aroused a stormy controversy, the echoes of which have still not wholly died down. At a time and in a country when the material sciences were achieving such remarkable successes and when materialism was in consequence becoming almost an article of faith, Fechner's philosophical speculations¹ (which in point of fact are tactfully played down in the *Elemente*) could hope for nothing but incredulity or scorn. On the other hand, his scientific contributions, as Wundt himself records, "immediately excited widespread interest and attention"². In the very year in which the book was published, Helmholtz in his own monumental work on *Physiologische Optik* cited Fechner's results with high praise. Mach at once embarked on a series of experiments to determine whether the law held good for the perception of time. On the other hand, younger psychologists, like Müller, Delboeuf, and Stumpf, found much to criticize and correct. Nevertheless, even they based their contentions on researches planned along the novel lines which Fechner himself had laid down for psychological investigations.

Fechner now turned his mind to the conquest of fresh fields. Having spent very nearly ten years on his psychophysical studies, he devoted yet another decade to the foundation of experimental aesthetics. As Woodworth remarks, at first sight this sudden change of interest appears "somewhat surprising". Yet, when we grasp the underlying philosophical theory that prompted the various types of question that Fechner was ceaselessly raising, the relationship becomes plain. Convinced as he was of the objective reality of the world of spiritual values, he now wanted to examine the aesthetic aspect as he had already examined the intellectual. As before, he proposed to investigate it, not 'from above' by the high *a priori* approach, but once again by

¹ Today an attempt to construct a 'world-model' in terms of spherical geometry no longer looks quite so fanciful as it did in Fechner's time. Eddington suggested that his 'uranoid' might have both psychical and physical aspects. The 'O-structure' set forth by C. C. L. Gregory and Anita Kohsen in their *Introduction to Psychophysical Cosmology* (1959), with its free use of logarithmic equivalences, is very like Fechner's theory in modern dress. His double-aspect hypothesis is in many ways similar to the theories of 'neutral monism' as developed by Mach, James, and Bertrand Russell, though the details vary widely. In some respects perhaps the views of Münsterberg come closer still, since in his philosophical writings he endeavoured to combine a parallelistic theory with a strong emphasis on the objectivity of values. As an interpretation of the relation between body and mind psychophysical parallelism became the fashionable theory till Ward, James, and McDougall revived the hypothesis of interaction. But the parallelist too often forgets that, in the relevant field of sense-perception, Fechner had demonstrated that the 'ordo et connexio idearum' did *not* (as Spinoza held) exhibit a simple one-to-one correspondence with the 'ordo et connexio rerum'. Nowadays, however, most psychologists would probably agree that the difference (which Fechner sought to express by substituting a logarithmic function for sheer equality) is the effect of biological requirements, acting through natural selection, rather than the reflection of some universal metaphysical relationship.

² W. Wundt, *G. T. Fechner: Rede zur Feier seines 100-jährigen Geburtstages* (1901).

trying to discover the coordinating laws by patient observation and experiment. He began by measuring cards, crosses, pictures, doors, windows, and the characteristic features of famous buildings in order to determine what kinds of proportion (if any) the artist, the architect, and the voice of popular preference tended unconsciously to select. This was followed by methodically planned experiments with rectangles, etc., of varying proportions to determine how far people preferred those conforming to the canon of the 'golden cut'. As before, his methods and conclusions, like those of most pioneers in any new sphere, proved to be no more than the first tentative approximations. But once again he had broken new ground for scientific study by the more exact procedures of experiment and statistics; and had successfully applied the methods of experiment and mathematical analysis to questions which had hitherto been answered solely on the basis of impression, authoritative dogma or speculative theory¹.

In 1879, the year in which Wundt opened at Leipzig the first psychological laboratory, Fechner returned to his metaphysical speculations with a work entitled *Die Tagesansicht gegenüber der Nachtansicht*. The universe is depicted as a living and conscious organism, with articulate parts, each of the parts, the stars, the planets, the earth, as well as plants, animals, and human beings, having its own organization—an organization which implies life and therefore in some sense a soul². The remainder of his publications were chiefly concerned with replies to criticisms of his psychological researches and several important amplifications or revisions. The last of these (except one or two published posthumously) was a long article on 'Psychische Massprincipien' which appeared in 1887. And in the same year, at the advanced age of 86, he died.

¹ The references to Fechner's work on aesthetics which the ordinary student usually encounters (e.g., the excellent chapter on 'Experimental Aesthetics' in Woodworth's *Experimental Psychology*) hardly make clear the historical background of the theory which Fechner had in mind. The classical principles of design had been systematically laid down by the Roman architect Vitruvius in a treatise dedicated to the Emperor Augustus—a work which was still cited as authoritative even as late as the time of the *Aufklärung*. The main theme was that beauty of every type (including the human form) was based on the relations of the parts to the whole: the key-words were *ordinatio*, *dispositio*, *symmetria*, *et proportio*, each formally defined: "pulcritudo enim ex ordinatione quadam constat quae graece *taxis* dicitur" (*De Architectura*, esp. *Lib. I*, ii and III, i). During the Renaissance "the essence of architecture came once again to be conceived as essentially a matter of mathematics and philosophy". The doctrine is perhaps most succinctly summed up by Palladio (1570): "The beauty of an edifice consists in an exact proportion of the parts within themselves and of each part with the whole: for a fine building ought to appear as an entire and perfect body wherein every member agrees with its fellow and each so well with the whole that it may seem absolutely necessary to the being of the same". The Romantic movement in poetry, art, music, and architecture led, about the middle of the nineteenth century, to violent controversies between critics of the opposing schools. In considering Fechner's selection of buildings, etc., it should be remembered that (except in the much-debated completion of Cologne cathedral) the Gothic revival never achieved the same popularity in Germany that it gained over here.

² The book displays, as Murphy observes, "an insistence on the meaning and value of all the known universe expressed with a lyrical beauty which moved even William James to an outburst of enthusiastic welcome to a kindred spirit" (*Historical Introduction to Modern Psychology*, p. 91; the reference is to James' *Pluralistic Universe*, 1909., chap. iv).

The far-reaching influence which Fechner's *Elemente* exerted on the young science of psychology during the next half century may be gathered from a glance at the textbooks still published towards the close of that period. In the second edition of Külpe's *Outlines of Psychology* (1901) well over one quarter of the volume is devoted to methods and results initiated by Fechner or modelled on his work. And nearly half of Brown and Thomson's standard book on *Mental Measurement* is concerned with 'Psychophysics'. Fifty years ago almost every student of psychology in this country began his laboratory course with detailed exercises on the psychophysical methods, usually taken from Titchener's four-volume manual of *Experimental Psychology*. As a training this was admirable ; but as a first introduction to psychology as a natural science it was decidedly repellent. With the increasing extension of experimental and statistical procedures to the wider and higher levels, psychophysics were gradually squeezed out ; and in many contemporary textbooks of psychology Fechner's name and work are not even mentioned.

Meanwhile, however, both in spirit and in detail Fechner's notion of combining experiment with statistics had spread from general psychology into more practical fields. Galton made a thorough study of Fechner's *Elemente* ; and in his own investigations of individual differences proposed "the use of tests, which, objectively speaking, run in a geometric series, but, subjectively, in an arithmetical series". In Whipple's *Manual of Mental and Physical Tests* most of the procedures for determining 'sensory capacity' are derived from Fechner's work and employ his terminology. However, Galton himself preferred to regard 'psychometry' as merely one branch of the wider science of 'anthropometry' ; and accordingly based the greater part of his statistical inquiries on the older work of Quetelet in the field of physical anthropology. The attempts to measure 'higher mental processes', which he subsequently devised or suggested, called for a rather different approach ; and Galton accordingly developed in a rough but practical form the novel methods of ranking and rating.

Those of us who, in the early days of mental testing, took our notions of the problem direct from Galton or Karl Pearson, and (like Brown, Flugel, and myself) had already received a thorough grounding in the psychophysical methods, naturally adapted the principles of the latter to solve the former. There is, as I ventured to describe it, a kind of "reciprocity between difficulty of test and capacity of testee : in measuring a man's sensory capacity we determine the weakest stimulus he can correctly detect in (say) 50 per cent of the trials ; in scaling tests we determine the difficulty of each item or test-question by specifying the ability required to perform it, e.g., the youngest age at which 50 per cent of the age-group can solve it correctly"¹.

¹ It was on this basis I suggested standardizing the items in the Binet-Simon Age-Scale. "To be assigned to any given year a test should be passed by approximately 50 per cent of the children who are nominally of a year below" (*Mental and Scholastic Tests*, 1921, p. 140). And it was along similar Fechnerian lines that I sought to determine the differences in difficulty between

From educational psychology the basic notions quickly spread to the fields of occupational psychology and industry. The problems of vocational guidance and selection often involve the testing of sensory capacities—particularly those of vision, hearing, touch, and smell—and can readily be measured in terms of ‘thresholds’. Psychologists and industrial engineers have successfully adapted psychophysical techniques to the study of illumination, ventilation, and noise. Still more recent applications have carried the same general principles beyond the limits of psychology: they have, for example, with minor modifications, been usefully employed in the assessment of the toxic capacities of drugs and of the fulminating point of explosives¹.

There can therefore be little doubt that both psychology and statistics owe far more to Fechner than is commonly recognized. Unfortunately the prestige of William James, and his amusing quips about ‘Old Caspar’ and ‘little Peterkin’, still carry greater weight than the patient study of the historical facts. Indeed it is difficult to avoid the suspicion that those who still quote them betray their contempt more plainly than their familiarity.

In what then does our debt consist? So far as the *Elemente* are concerned (and that is our primary interest here) we get first of all the systematic formulation of the three main psychological methods². These, as Boring says, “have stood the test of time; they are still used with minor modifications in the

the various test-items (*ibid.*, pp. 138f. For a fuller discussion see ‘Test-Construction and the Scaling of Items’, this *Journal*, IV, pp. 95–129.)

In America, on the other hand, so Guilford declares, mental testing and psychophysical experimentation developed in relative independence: “each has gone its own way, relatively unaware of the fundamental unity of all methods of psychological measurement” (*Psychometric Methods*, 1936, p. 9). Nevertheless, Thorndike and Thurstone undoubtedly appreciated the implications of psychophysical principles for mental testing; and Guilford has repeatedly emphasized the value of ‘rapprochement’ between the two.

¹ Thus, in studying the toxicity of a chemical substance the lethal dose is defined in terms of a ‘threshold’, i.e., as the minimum amount found to kill 50 per cent of the organisms to be treated with it. The equivalent of Fechner’s logarithmic law is then assumed to hold good of the relation between the amount of the substance and its effect on mortality. Investigators who have adopted the method commonly substitute the principle of maximum likelihood for that of least squares: this has the advantage of yielding better estimates of the statistical significance of the values obtained and of their ‘confidence limits’. The final solution entails the use of weights (Fisher and Yates, *Statistical Tables*, pp. 8–11) which are proportional to what the psychologist knows as the Müller-Urban weights. The whole procedure as thus modified has been termed the ‘probit method’; but in practice (as actual comparisons have shown) the essential results are virtually the same as those reached by Fechner’s ‘constant method’: (cf. this *Journal*, IV, 1951, p. 114 and refs., also R. A. Fisher, *Design of Experiments*, pp. 236f.).

² Fechner with his usual modesty gives full credit to astronomers and physiologists who had occasionally used (in a much cruder form) one or other of the procedures. But it was he who turned them into efficient instruments of psychological research. And in the course of the work he introduced many novel and serviceable devices—e.g., the ‘differential threshold’, the ‘median’ as the ideal method of determining a ‘central tendency’ in skewed distributions, and the application of the normal curve (familiar enough to astronomers and physicists) for describing data obtained by the psychologist. Moreover, his formula anticipated and encouraged the free use of logarithmic relations and logarithmic transformations for problems that would otherwise prove intractable.

laboratory today, and have proved applicable to problems of which Fechner himself never dreamt". Secondly, by his own example Fechner provided a model for the study of sensory phenomena with the aid of exact experimental techniques—a model which served, not only for determining the range and discriminative capacities of the human mind in general (and particularly their limitations), but also, in the hands of his later followers, for assessing the sensory capacities of different individuals. Thirdly, whatever we may think of the famous 'law', as he interpreted it, his demonstrations certainly brought home to the psychologists of his day the idea that they should henceforth regard the subject of their study as a genuine branch of natural science, involving the systematic search for general principles or 'laws' by the methods adopted in other sciences.

The methods, the results, and the law—all since revised or modified in consequence of later research—these are the main contributions which earlier writers discerned in the two volumes of the *Elemente*. But, looking back after a century has passed, it seems plain that we owe him far more than this. Fechner's true claim to greatness derives, not from his specific 'psychophysical methods' or from his cherished 'psychophysical law', but from the fact that he was the first to suggest the idea of mental measurement as a basic technique, applicable to the whole field of psychology, and at the same time to recognize that, in psychology in contrast to the more elementary sciences, this would entail the systematic use of statistical procedures. Before 1860 the idea of consistently applying experimental methods and numerical analysis to solve the problems of the psychologist was almost undreamt of; after 1860 they became an essential part of the science. And yet Fechner himself, in spite of his insistence on the needs for a quantitative treatment, never loses sight of the qualitative aspects. On all these grounds therefore he fully deserves a high and honourable place in the history of science as "one of the four or five great founders of modern psychology".

CYRIL BURT

APPROXIMATE METHODS IN FACTOR ANALYSIS

By D. N. LAWLEY

Mathematical Institute, University of Edinburgh

The following paper deals with the problem of estimating a factor matrix, when certain specified factor loadings are known or assumed in advance to be zero. Two simple procedures are proposed, the first for uncorrelated factors, the second for correlated factors. An approximate test for the significance of the residuals is also described, and the whole process illustrated by a worked example.

I. INTRODUCTION

In a previous paper (Lawley [1]) I considered the estimation by maximum likelihood of factor loadings when a certain number, in specified positions, are assumed *a priori* to be zero. The process of estimation is iterative, and requires an initial set of trial loadings. One way of providing them is to carry out a centroid analysis with subsequent rotation. However, although the centroid method is useful where little previous knowledge is available and where the factor analysis is mainly exploratory, it is somewhat wasteful of time and effort in cases where a sufficient number of loadings can be postulated to be zero in advance. If the zero loadings are such as to determine the factors uniquely it is possible, by processes such as those described below, to arrive at a set of approximate loadings which does not require rotation. Even if there is no intention of using maximum likelihood estimation, the methods should still be of some value, since the estimated loadings can be made the basis of an approximate χ^2 test for the significance of the residuals. Two methods will be given, one for orthogonal factors, and the other for oblique or correlated factors.

II. A METHOD FOR ORTHOGONAL FACTORS

The method described for orthogonal factors is to a certain extent arbitrary, since the values of the loadings will depend to a slight extent on the order in which the factors are considered. The general principle is that the factors are ordered roughly according to the number of zero loadings which they have, those with the highest number being considered first. For the method to work it is necessary that the r -th factor in this order should have at least $k-r$ zero loadings, where k is the total number of common factors assumed to be present. The k -th factor need have no zero loadings, and could thus correspond to a general factor.

By way of illustration we shall use the data on which maximum likelihood estimation was performed in the paper referred to above. With eight test

variates the standardized sample covariance matrix A was as follows (elements below the diagonal being omitted) :

1.000	0.691	0.679	0.149	0.409	0.382	0.346	0.270
	1.000	0.791	0.285	0.505	0.443	0.471	0.355
		1.000	0.314	0.472	0.372	0.426	0.254
			1.000	0.470	0.385	0.434	0.218
				1.000	0.680	0.639	0.504
					1.000	0.576	0.452
						1.000	0.395
							1.000

We first guess or estimate the communalities. In actual fact estimates previously obtained will be used ; but this is of no moment except that it facilitates comparison with the maximum likelihood method. The communalities, in order, are as follows : 0.638 0.782 0.843 0.440 0.757 0.619 0.561 0.368. Let A_0 be the matrix formed from A by replacing the unities along the diagonal by these values.

The hypothesis with which we start is that there are three common factors, and that the required pattern of loadings (with x 's representing those which are not necessarily zero) is as follows :

x	x	x	x	x	x	x	x
x	x	x	0	0	0	0	0
0	x	x	x	x	x	x	0

Corresponding to these three factors we write down a matrix Q' , with three rows, by which A_0 is to be pre-multiplied. The only essential condition which Q' must satisfy is that its rows should be linearly independent. The actual form of Q' will affect merely the efficiency of the estimation process. A simple general rule for choosing Q' is to put zeros for elements corresponding in position to zero loadings and unities elsewhere. Thus in our example we shall take Q' to be

1	1	1	1	1	1	1	1
1	1	1	0	0	0	0	0
0	1	1	1	1	1	1	0

In certain cases this rule would have to be slightly modified, since, if strictly applied, it might give rows which were linearly dependent. Thus, for example, the first row might be merely the sum of the second and third rows. Simple modifications of the rule can, however, easily be devised to cope with this difficulty.

The matrix $Q'A_0$, whose elements are merely sums or partial sums of the columns of A_0 , is easily found to be

3.564	4.323	4.151	2.695	4.436	3.909	3.848	2.816
2.008	2.264	2.313	0.748	1.386	1.197	1.243	0.879
2.656	3.277	3.218	2.328	3.523	3.075	3.107	2.178

We also calculate the matrix

$$H = Q'A_0Q = \begin{bmatrix} 29.742 & 12.038 & 23.362 \\ 12.038 & 6.585 & 9.151 \\ 23.362 & 9.151 & 18.528 \end{bmatrix},$$

where Q is the transpose of Q' .

The next stage of the process is to determine three row vectors z'_1 , z'_2 , and z'_3 by which $Q'A_0$ is to be pre-multiplied. The elements of z'_r , $Q'A_0$ are to be proportional to the estimates of the loadings in the r -th factor, and (for $r < k$) we must choose z'_r such that certain of these elements are as small as possible. In order that the factors should be orthogonal z_1 , z_2 and z_3 must also satisfy the orthogonality relations

$$z'_1 H z_2 = z'_2 H z_3 = z'_3 H z_1 = 0.$$

In our example we start with the second factor, since this has the greatest number of zero loadings. We must choose z'_2 so that the last five elements of $z'_2 Q'A_0$ are small. This could be done by trial and error; but we shall employ the method of least squares, and minimize the sum of squares of these elements. Let us write $z'_2 = [x_2 \ 1 \ y_2]$. Then we choose x_2 and y_2 so as to minimize the expression

$$(2.695 x_2 + 0.748 + 2.328 y_2)^2 + \dots + (2.816 x_2 + 0.879 + 2.178 y_2)^2.$$

This leads to the "normal" equations

$$\begin{aligned} 64.958 x_2 + 52.011 y_2 + 20.102 &= 0, \\ 52.011 x_2 + 41.684 y_2 + 16.082 &= 0. \end{aligned}$$

Hence

$$z'_2 = [-0.5812 \quad 1.0000 \quad 0.3394],$$

and

$$z'_2 H = [2.6810 \quad 2.6944 \quad 1.8614].$$

Taking next the third factor, we choose $z'_3 = [x_3 \ y_3 \ 1]$ in such a way that the sum of squares of the first and last elements of $z'_3 Q'A_0$ is minimized with z_3 satisfying $z'_2 H z_3 = 0$. Thus x_3 and y_3 must be chosen so as to minimize

$$(3.564 x_3 + 2.008 y_3 + 2.656)^2 + (2.816 x_3 + 0.879 y_3 + 2.178)^2,$$

subject to the condition

$$2.6810 x_3 + 2.6944 y_3 + 1.8614 = 0.$$

This gives

$$z'_3 = [-0.8095 \quad 0.1148 \quad 1.0000],$$

and

$$z'_3 H = [0.6678 \quad 0.1622 \quad 0.6670].$$

Finally, $z'_1 = [1 \ x_1 \ y_1]$ is determined by the conditions

$$z'_2 H z_1 = z'_3 H z_1 = 0,$$

i.e.

$$\begin{aligned} 2.6944 x_1 + 1.8614 y_1 + 2.6810 &= 0, \\ 0.1622 x_1 + 0.6670 y_1 + 0.6678 &= 0. \end{aligned}$$

Thus

$$z'_1 = [1.0000 \quad -0.3646 \quad -0.9125],$$

and

$$z'_1 H = [4.0351 \quad 1.2868 \quad 3.1187].$$

We now calculate $z'_1 H z_1 = 0.72012$, $z'_2 H z_2 = 1.76796$, $z'_3 H z_3 = 0.14505$, and divide the vectors z'_1 , z'_2 , z'_3 respectively by the square roots of these three values to obtain standardized vectors w'_1 , w'_2 , and w'_3 . This standardization is necessary in order that the factors should have unit variances. We find that

$$\begin{aligned} w'_1 &= [1.1784 \quad -0.4296 \quad -1.0753], \\ w'_2 &= [-0.4371 \quad 0.7521 \quad 0.2553], \\ w'_3 &= [-2.1256 \quad 0.3014 \quad 2.6258]. \end{aligned}$$

Pre-multiplication of $Q'A_0$ successively by w'_1 , w'_2 and w'_3 gives the estimated loadings for the three factors. The matrix L' of loadings is as follows :

$$\begin{array}{cccccccc} 0.481 & 0.598 & 0.438 & 0.351 & 0.844 & 0.786 & 0.660 & 0.599 \\ 0.630 & 0.650 & 0.747 & (-0.021) & (0.003) & (-0.023) & (0.046) & (-0.014) \\ (0.004) & 0.098 & 0.324 & 0.610 & 0.239 & 0.126 & 0.354 & (-0.002) \end{array}$$

The above values are, on the whole, not very different from the maximum likelihood estimates. Those in brackets are satisfactorily small ; we replace them by zeros.

New estimates of the communalities and of the residual variances may now be found in the usual way. The residual variances form a diagonal matrix V whose elements are the diagonal elements of $A - LL'$ and are, in order :

$$0.3717 \quad 0.2103 \quad 0.1452 \quad 0.5047 \quad 0.2305 \quad 0.3663 \quad 0.4391 \quad 0.6412.$$

III. A METHOD FOR CORRELATED FACTORS

With oblique or correlated factors the process of estimation is similar. Now, however, the order in which the factors are considered is irrelevant. For the factors to be determined uniquely each must have at least $k - 1$ zero loadings. The vectors z' , no longer have to satisfy orthogonality conditions, but instead it is necessary to estimate the correlation coefficients between the factors. To ensure uniqueness in our numerical example we must make an alteration in the hypothesis regarding the factors. The pattern of loadings is the same as before except that the third and fourth x 's in the first row are replaced by zeros. The matrix Q' we now take as :

$$\begin{array}{cccccccc} 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{array}$$

Hence $Q'A_0$ is as follows :

$$\begin{array}{cccccccc} 2.736 & 3.247 & 2.994 & 1.941 & 3.494 & 3.152 & 2.988 & 2.344 \\ 2.008 & 2.264 & 2.313 & 0.748 & 1.386 & 1.197 & 1.243 & 0.879 \\ 2.656 & 3.277 & 3.218 & 2.328 & 3.523 & 3.075 & 3.107 & 2.178 \end{array}$$

We also find

$$H = Q'A_0Q = \begin{bmatrix} 17.961 & 8.977 & 17.816 \\ 8.977 & 6.585 & 9.151 \\ 17.816 & 9.151 & 18.528 \end{bmatrix}$$

In this case we can choose the vector $z'_1 = [1 \ x_1 \ y_1]$ so that the loadings of the third and fourth variates in the first factor are exactly zero. The equations for x_1 and y_1 are

$$\begin{aligned} 2.313 x_1 + 3.218 y_1 + 2.994 &= 0, \\ 0.748 x_1 + 2.328 y_1 + 1.941 &= 0. \end{aligned}$$

Hence

$$z'_1 = [1.0000 \quad -0.2431 \quad -0.7556].$$

Similarly we can choose $z'_3 = [x_3 \ y_3 \ 1]$ so that the loadings of the first and last variates in the third factor are exactly zero. The equations for x_3 and y_3 are

$$\begin{aligned} 2.736 x_3 + 2.008 y_3 + 2.656 &= 0, \\ 2.344 x_3 + 0.879 y_3 + 2.178 &= 0; \end{aligned}$$

and thus

$$z'_3 = [-0.8857 \quad -0.1159 \quad 1.0000].$$

For the second factor we choose $z'_2 = [x_2 \ 1 \ y_2]$ in such a way that the sum of squares of the last five elements of $z'_2Q'A_0$ is a minimum. The normal equations for x_2 and y_2 are

$$\begin{aligned} 40.333 x_2 + 40.909 y_2 + 15.842 &= 0, \\ 40.909 x_2 + 41.684 y_2 + 16.082 &= 0, \end{aligned}$$

giving

$$z'_2 = [-0.3196 \quad 1.0000 \quad -0.0721].$$

Let Z' be the matrix whose three rows are, in order, z'_1 , z'_2 and z'_3 . Then, finding $Z'HZ$ and its inverse, we have

$$Z'HZ = \begin{bmatrix} 1.0020 & -0.3936 & -0.5140 \\ -0.3936 & 2.2794 & 0.0380 \\ -0.5140 & 0.0380 & 0.8688 \end{bmatrix}, \quad (Z'HZ)^{-1} = \begin{bmatrix} 1.5688 & 0.2556 & 0.9169 \\ 0.2556 & 0.4807 & 0.1302 \\ 0.9169 & 0.1302 & 1.6878 \end{bmatrix}.$$

We put the latter matrix in the form DPD where D is a diagonal matrix with elements (1.2525 0.6933 1.2992) and where

$$P = \begin{bmatrix} 1.000 & 0.294 & 0.563 \\ 0.294 & 1.000 & 0.145 \\ 0.563 & 0.145 & 1.000 \end{bmatrix}.$$

The elements of P represent estimates of the factor correlation coefficients.

In order to obtain loadings in factors with unit variances we find

$$W' = DZ' = \begin{bmatrix} 1.2525 & -0.3045 & -0.9464 \\ -0.2216 & 0.6933 & -0.0500 \\ -1.1507 & -0.1506 & 1.2992 \end{bmatrix}.$$

Pre-multiplication of $Q'A_0$ by W' then gives the matrix L' of loadings as follows :

0.302	0.276	0	0	0.620	0.673	0.424	0.607
0.653	0.686	0.779	(-0.028)	(0.010)	(-0.022)	(0.044)	(-0.019)
0	0.180	0.387	0.678	0.348	0.188	0.411	0

The values obtained are again not very different from the maximum likelihood estimates. As before we replace the bracketed entries by zeros.

The diagonal elements of the matrix $A - LPL'$ provide estimates of the residual variances and constitute the diagonal matrix V . They are as follows :

0.3664	0.2178	0.1560	0.5403	0.2516	0.3693	0.4551	0.6316.
--------	--------	--------	--------	--------	--------	--------	---------

IV. EFFICIENCY OF ESTIMATION AND TESTS OF SIGNIFICANCE

It would be extremely difficult, if not impossible, to carry out a full investigation of the efficiency of the methods of estimation described above. In discussing the efficiency of the centroid method (Lawley [2]) it was supposed that the method was performed on an unstandardized covariance matrix and that the residual variances were known exactly. The latter condition is admittedly unrealistic, and would not be satisfied in practice except in artificial sampling experiments ; but a simplifying assumption of this kind seemed necessary before any progress could be made. Under these conditions it was possible to assess the efficiency of the centroid method and also to provide a criterion for testing the significance of the residual covariances. This criterion was distributed asymptotically (for large samples) as χ^2 with $\frac{1}{2}(p-k)(p-k+1)$ degrees of freedom, where p is the number of test variates and k the number of common factors. Under the same conditions these results could be used in the present case with the matrix Q playing precisely the same role as the Q previously used (where each element was either 1 or -1).

It is not easy to assess the effect of dispensing with these conditions and of taking account of sampling errors in the residual variances, particularly when inefficient estimates are used. The criterion for testing significance will now not follow an exact χ^2 distribution, even asymptotically. As a very rough approximation, however, we may still treat it as a χ^2 variate and reduce the number of degrees of freedom by p . The adequacy of this approximation will depend not only on the sample being a large one but also on p being sufficiently large compared with k to ensure that the reduced number of degrees of freedom is not too small. The inefficiency of estimation of the residual variances will become progressively less important as the number of degrees of freedom increases.

Instead of expressing the criterion in terms of residuals calculated in the usual way, we shall employ a method which is arithmetically more convenient. The method is the same whether the factors are orthogonal or correlated. In each case the number of degrees of freedom is taken as

$$\frac{1}{2}(p-k)(p-k+1) - p = \frac{1}{2}\{(p-k)^2 - (p+k)\}.$$

We find successively the matrices $L'V^{-1}$, $J=L'V^{-1}L$, $M'=J^{-1}(L'V^{-1}A-L')$ and $X=V^{-1}(A-LM'-V)$. The matrix $(A-LM'-V)$ is analogous to the usual residual matrix, but is not symmetrical and its diagonal elements are not necessarily zero. The χ^2 criterion is calculated as

$$\frac{1}{2}n \operatorname{tr}(X^2) = n \left\{ \frac{1}{2} \sum_i (x_{ii})^2 + \sum_{i < j} (x_{ij} x_{ji}) \right\},$$

where tr denotes trace or sum of diagonal elements and x_{ij} is the typical element of X .

The data in our example are based on a sample of 211 individuals. Hence $n=210$, $p=8$, $k=3$ and the number of degrees of freedom for χ^2 is 7. In the case of orthogonal factors the value of χ^2 was found to be 4.2, while for correlated factors it was 3.7. In either case it is well below expectation and would lead us to accept the hypothesis, though with only seven degrees of freedom the test of significance must be viewed with some suspicion.

In general, when the criterion is either well below or well above the chosen significance level, it should be safe to accept or reject the hypothesis as a result of this test. In borderline cases a test based on maximum likelihood estimation would be necessary to reach a firmer decision.

V. SUMMARY

Two simple methods are given for estimating factor loadings when a certain number of these, in specified positions, are assumed to be zero. One method is for orthogonal factors, the other for correlated factors. An approximate test of significance based on the estimated loadings is suggested.

REFERENCES

- [1] LAWLEY, D. N. (1958). Estimation in factor analysis under various initial assumptions. *Brit. J. Statist. Psychol.*, XI, 1.
- [2] LAWLEY, D. N. (1955). A statistical examination of the centroid method. *Proc. Roy. Soc. Edin.*, A, LXIV, 175.

A STUDY OF SUBJECTIVE PROBABILITY

By H. C. A. DALE

Medical Research Council, Applied Psychology Research Unit

Adult subjects predicted the way a small number of items would be selected by chance from a long list of items. The frequencies of their predictions were then compared with the objective probabilities. The results suggest that the subjects avoided what they thought to be unlikely configurations, but that, in doing so, they considered certain aspects of a selection only, and in these appeared to be guided by one or two simple rules which showed only a limited accord with the laws of chance.

I. INTRODUCTION

If a person is confronted with a row or column of similar things, and asked to pick out a few at random, how will he choose? Previous research suggests that he would not in point of fact pick them at random: in analogous situations people do not behave as generators of random sequences.

Relevant investigations are of two kinds. First, there are the two-choice serial guessing experiments in which subjects have to guess which of two alternatives will occur on each of a succession of trials (see Hake's review [1], Ross [2], Goodfellow [3]). Secondly, there is the study by Cohen and Hansel [4] in which subjects had to guess the colour and position of four beads, drawn at random from a bag containing blue and yellow beads and placed in a row of four beakers, one into each. Here the subject had not only to decide how given numbers of beads would be arranged, but also what colours would be drawn.

The two-choice guessing experiments show that in this task subjects generally tend to alternate: having guessed one alternative, they are more likely to choose the other. If we regard our column as a series scanned from top to bottom, we would expect on the basis of these results that subjects would not pick out items which were adjacent to each other. Cohen and Hansel's study, however, yielded results conflicting with this expectation. When their subjects predicted that one yellow and three blue beads would be drawn (or vice-versa) they tended to place the odd one in one of the end positions. When they predicted that two blue and two yellow would be drawn, they arranged them in the three possible ways (abab, aabb, abba) with equal frequency. Thus they showed no preference for alternation. In view of this conflict a direct study has been undertaken. Two techniques have been used: first, subjects have made free selections; secondly, a number of selections have been made by the experimenter and the subjects have been required to choose between them.

II. FREE SELECTION

Method. In the first investigation the subjects freely selected five items from a column of fifteen. The object of the investigation was disguised, and the randomness involved made implicit only. The subjects consisted of 44 young enlisted men, tested individually. They were given a cyclostyled sheet, headed 'Five Draws'. On this a list of fifteen pairs of Indian town names ran down the left-hand side labelling the 15 rows of a 15×20 matrix of cells. The subjects were told that this form was a Pakistani football coupon, and that the names were those of teams. Fifteen matches were to be played. The subjects were to pick five which they thought would result in a draw, and mark the cells of the first column accordingly. Having done so, they were then asked to make four entries on the same form.

Results. The results were analysed first for evidence that items in certain positions were selected more frequently than others, and secondly for evidence that certain configurations would be selected more frequently than others. Three aspects of the configurations have been considered (i) Range (i.e., the distance between the topmost item selected and the bottom-most) ; (ii) Symmetry ; (iii) Bunching, measured by the largest group of adjacent items selected. Each aspect will be discussed independently, although they are to some extent interdependent.

Three possible models are examined. Model I supposes that each subject chooses items at random, so that the choices of the group correspond to those of a random generator. Model II supposes that, when each subject makes a selection, he looks upon it as one of the total family of all possible selections, considers the three relevant aspects in this family of selections (i.e. Range, Symmetry, and Bunching), and then for each aspect considers each possible kind of configuration that can arise ; finally, having done this, he chooses between the possible kinds at random. Thus, according to this model, we would expect the group as a whole to choose the different kinds of configuration with equal frequency. To offer a concrete example : each subject would consider Symmetry, realize that the selection can be symmetrical or asymmetrical and choose between these two at random. Model III supposes that as in the case of Model II, each subject considers the different aspects and the different kinds of configuration, but, instead of choosing between them at random, he chooses that kind of selection most likely to occur by chance : thus, since asymmetrical selections are more likely to occur by chance than symmetrical, we should expect all the subjects to produce asymmetrical selections.

Preferences for Positions. From Table I it can be seen that the subjects tended to avoid the ends of the column and to prefer those items which were a little above the middle. Had they selected items at random, each would have been chosen with equal frequency. To test the significance of the departure of the observed distribution of choices from that expected on this random hypothesis a modification of the chi-square test due to Stone [5] has been

used. This is applied to the results of the first entry ; it is not possible to use the data from all five entries. Stone has shown that when each subject chooses a items from a total of c , then

$$\sum \frac{(Q - \bar{Q})^2}{\bar{Q}} \times \frac{c-1}{c-a}$$

(where Q is the number of times each item is chosen, and $\bar{Q}$ the mean number of choices per item) is distributed as chi-square with $c-1$ degrees of freedom. The value here obtained, 13.4, is not significant. Data from some ancillary observations in which 25 subjects picked 6 items from a column of 26, however, showed a similar deviation from true randomness (42.5 with 25 degrees of freedom, $p < 0.025$). For this supplementary experiment two new lists were used. The first was derived from the original list by adding 11 pairs ; the second by rearranging the order of the 26 pairs. Eleven subjects were given the first list ; the remaining 14 the second. The value of 42.5 was obtained from the total results. The effect was manifestly due to preferences for certain positions. Had the preferences been associated with names rather than with positions no such deviation would have appeared in the total results.

TABLE I. FREQUENCY WITH WHICH ITEMS AT EACH POSITION OF THE LIST WERE SELECTED

Position	Selection					Total
	1st	2nd	3rd	4th	5th	
1	11	16	15	14	13	69
2	16	20	15	15	14	80
3	18	17	14	16	17	82
4	17	17	12	17	15	78
5	17	19	14	13	23	86
6	14	19	15	16	13	77
7	16	14	16	18	19	83
8	13	15	9	13	20	70
9	18	14	22	19	11	84
10	15	11	16	13	18	73
11	13	19	15	15	14	76
12	17	11	19	12	10	69
13	8	11	11	15	11	56
14	10	10	16	10	13	59
15	17	7	11	14	9	58

Range. The details are summarized in Table II. Model I will fit these data reasonably well. When each entry is taken separately, no significant difference is found. When, however, the successive entries are compared with each other, a trend is seen to run through the whole series : the later entries include a steadily increasing number of configurations showing small ranges. Even so, however, a test of significance yields a negative result. Hence there is so far no ground for rejecting the model.

Model II entirely fails to fit the data. The subjects did not produce selections with each possible range showing equal frequency ; and the chi-square test indicates that the divergence is fully significant at the 1 per cent level. The range which would occur most commonly by chance is 13 ; and according to Model III all the subjects should have made one or more selections with this range. This was not the case. It may be noted, however, that the few deviations from Model I were caused mainly by the frequent occurrence of selections having this particular range.

Symmetry. If, when the column is folded so that the topmost item selected lies above the bottom-most, the configuration of items chosen forms two mirror halves, the selection may be termed 'symmetrical'. 6, 8, 5, 5, and 8 of the selections were symmetrical with each of the successive entries ; the number expected by chance is 1.3.

None of the models entirely fits these results. The majority of the subjects certainly chose that type of selection which would be most likely to occur by chance. Nevertheless, the deviation from true randomness was not in the direction which would be predicted on the basis of Model III. Models I and II also fail to fit ($p=0.05$ and 0.01 respectively).

TABLE II. SELECTIONS ANALYZED FOR RANGE

The figures in the body of the table show the number of subjects making selections having a given range ; the figures in the last row give the number expected in a random distribution.

	Range											
	5	6	7	8	9	10	11	12	13	14	15	Total
1st Entry	0	0	0	1	4	4	6	9	11	6	3	44
2nd Entry	0	1	0	3	5	6	5	6	14	2	2	44
3rd Entry	1	1	0	1	4	6	3	8	13	4	3	44
4th Entry	1	1	1	2	2	5	4	6	9	11	2	44
5th Entry	2	0	2	0	9	3	8	6	5	6	3	44
Mean	0.8	0.6	0.6	1.4	4.8	4.8	5.2	7.0	10.4	5.8	2.6	44
Number expected	0.2	0.6	1.3	2.3	3.6	4.9	6.1	7.0	7.3	6.5	4.2	44

Bunching. The term 'bunching' is used to designate the choice of adjacent items. When 5 items are selected from an array of 15, there are only 7 different possibilities as regards bunching : at one extreme all 5 items would be adjacent to each other, at the other each of the 5 would be isolated from the others. The complete set of possibilities is set out in Table III with the frequencies observed and those expected by chance.

Model I is here quite unsuitable. With every entry there were far too many selections in which all the items were isolated (the deviation was significant at the 1 per cent level). Model II must also be rejected, since the different kinds of bunching did not occur with equal frequency. According to Model III every subject should have made selections in which there was one pair of adjacent items. Rather more actually did so than would be expected had they all selected items at random. But this could hardly account for the main

deviation. Possibly the common occurrence of selections in which all the items were isolated resulted from the kind of behaviour predicted by the model combined with an ignorance of the laws of chance. To extend the model in this way, however, would allow it to fit any observation, and would thus render it meaningless. Once again a trend seems discernible, but it is not significant.

General Conclusions. It is evident therefore that none of the models fits the data particularly well. As we have seen, Model I must be rejected because of the position preferences, the symmetrical configurations, and the avoidance of bunching. Model II is wildly inappropriate. Model III deserves more consideration, since the deviations from randomness in both range and bunching are in the direction that it predicts : in both respects the kind of configuration most likely to occur by chance appeared with special frequency. The deviations from the model in regard to bunching suggests that, while behaving in the general way predicted, the subjects were not familiar with the strict laws of chance. In regard to symmetry, however, the deviation from randomness definitely contradicts the model, since fewer subjects chose the commonest type of selection that randomness would tend to produce. Hence, as it stands, this model must also be rejected. In the form proposed it involves some rather rigid assumptions, viz., that the subjects choose from all possible selections, that they consider the particular criteria we have enumerated, that they know which kind of selection is most likely to occur, and that all behave in the same way. It might therefore be worth examining an alternative with weaker assumptions, e.g., that *most* subjects behave in this general way, and that they consider only *some* of the aspects we have enumerated, and are in *some* cases ignorant of the true probabilities. However, once again such an alternative would fit almost any data.

The main assumption of Models II and III is that the subjects look on the product as a possible member of different classes of configurations, and then complete the selection so that it belongs to one class rather than another. So far, however, we have no evidence that complete selections belonging to one class are in fact preferred. It will be the purpose of the choice experiment described in the next section to seek this evidence.

Much of the observed divergence from true randomness could be accounted for by the relatively simple rule that subjects avoid distinctive items. End items are distinctive ; and the choice of an item within an array would also render the items adjacent to it distinctive. This rule is more parsimonious than Model III. It would account for the infrequency both of end cells and of bunching, and might possibly explain the preference for configurations with a range of 13, since the avoidance of bunching leads to the avoidance of small ranges while the avoidance of end cells makes the large values (14 and 15) unpopular. But since the rule has itself been suggested by the data, it would be illogical to claim the data support the rule. It should be possible to check its validity by further experiments in which items are rendered distinctive by some irrelevant cue.

Both Model III and this simple rule imply that the subjects adopted a process of elimination. This would be analogous to the procedure used by the statistician to produce a series of random numbers. He begins with a set of tracts, and rejects those which fail to satisfy the test for non-randomness. However, even the surviving tracts might still prove non-random if judged by a test he has not applied. This way of looking at the product of a human randomizer might be instructive, since we could perhaps describe his behaviour in terms of the tests he tacitly applies and fails to apply.

III. A CHOICE EXPERIMENT

In the next experiment the subjects had to choose between ready-made selections. A set of selections was prepared so that the three factors mentioned above—range, symmetry, and bunching—could be varied independently. As before, the task was presented in the form of a football-pool competition ; and each subject had to choose that member of each pair which he would prefer to enter for the competition. In preparing the list six cells were selected from a column of 26 in 30 different ways. The selections had five different kinds of bunching (L, K, J, H and D in Table IV), three different ranges (12, 18, or 26 cells), and were either perfectly symmetrical or asymmetrical. When the range was 12 or 18 the selection was positioned centrally. Each was printed by hand on a separate card. The increase in the number of arrays and choices was, it will be noted, necessitated by the requirements of the design.

Table IV summarizes the different sorts of bunching that may occur when 6 items are selected from 26, together with the chance probabilities. It will be seen that the size of the largest bunch does not uniquely determine the probability of the selection, though it gives a general indication. From the observations on free selection it might plausibly be predicted that the subjects would be guided almost entirely by the size of the biggest bunch. Accordingly, to test this prediction, three different comparisons were carried out.

The Experimental Design. The comparison of selections differing in range and symmetry was straightforward : there were three ranges ; half the selections were symmetrical, half asymmetrical. The comparison of bunching is more complex. In Table IV the kinds of bunching used are marked with an asterisk. If preferences for L selections are compared with those for H, J, K (taken as a group) and with D, the effect of the size of the biggest bunch can be readily seen : with L it is unity ; with H, J and K two ; with D four. If preferences for H, J and K are compared, any differences will show that some factor other than the size of the biggest bunch is operating. If preferences for H are compared with those for D, we have a case where the number of bunches is important objectively, and the selection having the largest single bunch (D) is more likely to occur by chance than the selection with the smallest bunch (H). With 30 different selections there are 870 pairs. This would be excessive for one subject, and so the task was shared. The plan of the experiment was

based on a Graeco-Latin square. Sixteen new subjects were used, and the 31 series given in the square allocated so that 15 of them each had two picked at random, while the sixteenth had the remaining one.

TABLE III. SELECTIONS ANALYSED FOR BUNCHING
Number of Subjects Producing Each Kind of Selection

Types of Bunching	Selection					Mean	Expected Frequency
	1st	2nd	3rd	4th	5th		
(5)	0	0	1	1	2	0.8	0.16
(4) + (1)	0	0	1	0	1	0.4	1.61
(3) + (2)	0	1	1	1	0	0.6	1.61
(3) + (1) ²	1	1	3	1	7	2.6	7.25
(2) ² + (1)	0	3	1	2	0	1.2	7.25
(2) + (1) ³	22	19	12	14	10	15.4	19.34
(1) ⁵	21	20	25	25	24	23.0	6.77

Note.—In the first column the figures inside the brackets indicate the number of adjacent cells, and the suffix indicates the number of times a bunch of that size occurs in the selections.

TABLE IV. THE PROBABILITY OF SELECTIONS SHOWING DIFFERENT TYPES OF BUNCHING

Code Letter	Number in Biggest Bunch	Type of Bunching	Probability	Summed Probabilities
A	6	(6)	0.00009	0.00009
B	5	(5)+(1)	0.00182	0.00182
C	4	(4)+(2)	0.00182	} 0.01915
D	4	(4)+(1) ^{2*}	0.01733	
E	3	(3) ²	0.00091	} 0.13955
F	3	(3)+(2)+(1)	0.03466	
G	3	(3)+(1) ³	0.10398	} 0.59881
H	2	(2) ^{3*}	0.00091	
J	2	(2) ² +(1) ^{2*}	0.15597	} 0.23570
K	2	(2)+(1) ^{4*}	0.44193	
L	1	(1) ^{6*}	0.23570	

Note. The figures for the probabilities are calculated by Mood's formula [6] from the random distribution of runs. The summed probabilities are based on the size of the biggest bunch. The asterisks in Col. 3 indicate the types of selection used in experiment II.

As an introductory task the subjects made some free selections much as in the first experiment, except that now they had to select 6 cells out of 26. After making 5 selections, they were told that some selections had been made with a random procedure by a friend of the experimenter. These were then shown in the order required by the Graeco-Latin square. The subjects were forced to choose from every pair the item which they preferred.

Results.—Table V shows the number of times each arrangement was preferred. As regards bunching, their preferences appeared to depend on the size of the largest bunch. A non-parametric statistic due to Friedman [7]

TABLE V. NUMBER OF TIMES EACH CONFIGURATION WAS PREFERRED COMPARED WITH THE EXPECTED NUMBERS

The numbers observed are shown in the S (symmetrical) and AS (asymmetrical) columns ; their sum under S+AS ; the sum expected, had choices been based upon relative probabilities for Range and Bunching, under E(S+AS).

Code Letter		L	K		J		H		D		Total	Total										
Type of Bunching Range S	AS	(1) ^a	(2) + (1) ^a		(2) ^a + (1) ^a		(2) ^a		(4) + (1) ^a		(S + AS)	E(S + AS)										
		S + AS	E(S + AS)	S	AS	S + AS	E(S + AS)	S	AS	(S + AS)	E(S + AS)											
12	34	68	13.19	30	27	57	50.58	33	35	68	46.64	24	23	47	17.53	12	20	32	35.09	272	163.03	
18	36	40	76	80.18	38	36	74	92.72	34	32	66	74.00	27	32	59	24.24	23	23	46	43.17	321	314.31
26	34	29	63	107.63	28	36	64	109.58	25	32	57	89.31	28	25	53	33.30	22	18	40	52.85	277	392.67
Sum	—	207	201.00	—	—	195	252.88	—	191	209.95	—	159	75.07	—	118	131.11	870	870.01				

indicated that the overall differences were significant ($0.05 > p > 0.01$). When the results were analyzed to make the three comparisons proposed above, it appeared (a) that the preferences for selections having different numbers of items in the largest bunch differed significantly ($p < 0.01$); (b) that there was no significant differentiation between the different selections which had the same number of items in the biggest bunch, viz. $(2) + (1)^4$, $(2)^2 + (1)^2$, and $(2)^3$; (c) that the arrangement $(2)^3$ was preferred to $(4) + (1)^2$ ($p < 0.025$). As regards range, selections with the middle value (18) were chosen most frequently, and those having large or small values with smaller but roughly equal frequencies. The overall differences were significant ($p < 0.05$). Symmetry proved to be of no account.

Had the subjects chosen according to the relative probability that certain kinds of configuration would occur by chance, we should have expected them to choose asymmetrical selections more frequently than symmetrical. They did not. It is still possible, however, that they selected in this way in regard to bunching and range. The number of choices then expected is shown in the E (S+AS) columns of Table V. Comparison with the results (S+AS) columns shows that, although there is some correspondence, this assumption does not fit the data very well ($\tau = 0.25$, which is not significant).

General Conclusions. The main fact which emerges from this further experiment is that, since the overall results show that some configurations were preferred to others, some kind of stereotype must exist for making this type of judgement, at any rate among the subjects we used. It follows that they could not have based their choices upon the true probabilities; for, had they done so, they would have had no reason to prefer one member of a pair to the other. The results suggest that they regarded a configuration as representative of a class of selections rather than as a particular instance, and therefore argued quite falsely: "This kind of configuration is less likely to occur by chance than that kind; therefore I will choose that one." If so, they displayed a second error, for the frequency with which the various configurations were chosen did not correspond to the frequency with which the various kinds would occur by chance.

The three comparisons between selections with different types of bunching demonstrated plainly that the subjects considered the size of the largest bunch only, and that their preferences depended on the number of items it contained, regardless of the number of bunches of a given size. The computations used for determining the objective probability took account, not only of this second factor, but also of the fact that, with 6 items and 26 cells, a little room is allowed for manoeuvre. There are comparatively few ways of arranging the six so that no two fall adjacently; yet the subjects preferred selections with six separate items to selections in which a pair was bunched—a preference which implies that they failed to realize that such patterns were unlikely to occur by chance. Apparently they adopted the rule that items are unlikely to be chosen from adjacent cells. Had a smaller number of items (or a longer list) been available,

the choices might have corresponded with those based on the chance frequencies. Knowing the results of this experiment, we can now suggest an alternative to Model III, namely, that the subjects consider certain aspects of the complete configurations which they might select, and from the entire family choose the kind which they think most likely to arise by chance. The results thus provide an index of what they apparently think most likely, at least among a small number of possibilities. Since the configurations preferred are of the sort produced relatively frequently, such a model seems reasonable.

The rule suggested by the analysis of the free selections, namely, that the subjects avoid distinctive items, would also account for the results of the experiment. It cannot be applied very rigorously here, since it describes the procedure used while picking *successive* items within a selection. Nevertheless, the configuration preferred would be of much the same kind as that which was preferred in this last experiment : configurations which contain bunching and configurations in which the end items are conspicuous would be avoided, and the subject would also be indifferent to symmetry. The only result which does not agree with expectation is the marked avoidance of configurations showing a smaller range.

IV. SUMMARY AND FINAL CONCLUSIONS

The foregoing investigations indicate, though doubtless with unavoidable limitations, what is the general nature of the expectations entertained by subjects like our own in regard to random events. They imply that subjective ideas of randomness are grossly erroneous—an outcome which is not very surprising since previous investigators have come to much the same conclusion. Many of the results obtained in our experiments might almost have been predicted from what is known about serial two-choice guessing behaviour, since the avoidance of bunching appears to correspond to the avoidance of runs on one particular alternative. But there are certain phenomena—for instance the avoidance of end items and the indifference to symmetry—which could not have been predicted in this way ; and indeed the findings of Cohen and Hansel are in actual conflict with our results. This conflict calls for explanation. There are obvious differences in the tasks imposed, and marked differences in the age of the subjects. Any of these factors might have led to the differences between the results ; and it remains to determine which were the most influential.

Further studies of this kind need to be undertaken before we can reach a comprehensive picture of subjective ideas of randomness. Nevertheless, it may be of interest to consider briefly what may be the reasons for the false ideas which are thus revealed. The explanations put forward in this paper suggest two possibilities. The avoidance of distinctive items implies that the subjects were trying to work from first principles, but in the end produced results which differed from true randomness because of what might be called a perceptual error. They argue, quite correctly, that certain items are unlikely to occur by chance ; but they fail to see that the argument might equally well

be applied to each and every item. Whitfield has written : " The usual human idea of randomness of a series of events is one which shows no apparent cause : *this is primarily a perceptual problem* " [8 ; my italics]. The second possibility is that the subjects based their judgement upon certain of their previous observations. If this were so, then either they have observed falsely—and there is evidence from other sources that persons do err when reporting the frequencies of random events which they claim to have observed [9]—or alternatively they have generalized from other situations, and their generalizations are inaccurate : the avoidance of bunching might plausibly be attributed to false generalization of this type¹.

REFERENCES

- [1] HAKE, H. W. (1955). The perception of frequency of occurrence and the development of 'expectancy' in human experimental subjects. *Ap. Information Theory in Psychology* (ed. Henry Quastler). Glencoe, Ill. : The Free Press.
- [2] ROSS, B. M. (1955). Randomization of a binary series. *Amer. J. Psychol.*, LXVIII, 136–138.
- [3] GOODFELLOW, L. D. (1938). A psychological interpretation of the results of the zenith radio experiments in telepathy. *J. Exp. Psychol.*, XXIII, 601–632.
- [4] COHEN, J. & HANSEL, C. E. M. (1955). The idea of a distribution. *Brit. J. Psychol.*, XLVI, 111–121.
- [5] STONE, M. (1960). An extension of the chi-squared test for randomness. *Brit. J. Statist. Psychol.*, XIII, pp 31 f.
- [6] MOOD, A. M. (1940). The distribution theory of runs. *Ann. Math. Statist.* X, 367–392.
- [7] FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.*, XXXIII, 675–701.
- [8] WHITFIELD, J. W. (1950). The imaginary questionnaire. *Quart. J. Exp. Psychol.*, II, 76–87.
- [9] PARDUCCI, A. (1957). Alternative measures for the discrimination of shift in reinforcement-ratio. *Amer. J. Psychol.*, LXX, 194–202.

¹The researches here described were carried out under the general direction of Dr. N. H. Mackworth. The writer wishes to thank Dr. M. Stone for statistical advice and the Royal Navy for supplying subjects.

AN EXTENSION OF THE CHI-SQUARED TEST FOR RANDOMNESS

By M. STONE

Medical Research Council Applied Psychology Research Unit

The following note describes a statistical test for random allocation, which has the property of being invariant even when the labelling of cells is changed.

Suppose that a objects are placed in c cells ($1 \leq a < c$) with the restriction that each cell can hold at most one object, and that data consisting of s independently generated choices are available. It is required to test whether the choices are consistent with random allocation, the alternative envisaged being that some cells will be preferred. Random allocation may be defined as one in which each of the $\binom{c}{a}$ possible choices has equal probability. Accordingly, let n_i denote the number of times the i 'th cell is chosen ($i = 1, \dots, c$). Then the test will be based on the statistic

$$X = \frac{(c-1)}{(c-a)} \sum_{i=1}^c \frac{(n_i - sa/c)^2}{sa/c}.$$

When $a = 1$, X is the ordinary chi-squared statistic. The following properties of X will now be established.

Under random allocation $E(X) = c - 1$ and $\text{var } X = 2(c-1)(1-1/s)$. Summations are from 1 to c except where otherwise indicated.

$$\begin{aligned} E(X) &= \frac{(c-1)}{(c-a)} \sum \frac{E(n_i - sa/c)^2}{sa/c} = \frac{(c-1)}{(c-a)} \sum \frac{s(a/c)(1-a/c)}{sa/c} = c-1. \\ E(X^2) &= \left(\frac{c-1}{c-a}\right)^2 \sum_i \sum_j \frac{E\{(n_i - sa/c)^2(n_j - sa/c)^2\}}{(sa/c)^2}. \end{aligned} \quad (1)$$

As a moment of the binomial distribution (Kendall [1], p. 117),

$$E(n_i - sa/c)^4 = 3s^2 \left(\frac{a}{c}\right)^2 \left(1 - \frac{a}{c}\right)^2 + \frac{sa}{c} \left(1 - \frac{a}{c}\right) \left[1 - \frac{6a}{c} \left(1 - \frac{a}{c}\right)\right] \quad (2)$$

for $i = 1, \dots, c$.

Now write $n_1 - sa/c = u_1 + u_2 + \dots + u_s$ and $n_2 - sa/c = v_1 + v_2 + \dots + v_s$ where $u_i = 1 - a/c$ if the first cell is in the i 'th choice
 $= -a/c$ otherwise

and similarly for v_i and the second cell;

then

$$\begin{aligned}
 E\{(n_1 - sa/c)^2(n_2 - sa/c)^2\} &= E \sum_{i=1}^s \sum_{j=1}^s \sum_{k=1}^s \sum_{l=1}^s u_i u_j v_k v_l \\
 &= E(\sum_{i=1}^s u_i^2 v_i^2) + E(\sum_{i \neq j} u_i^2 v_j^2) + 2E(\sum_{i \neq j} u_i u_j v_i v_j) \\
 &= s E(u_1^2 v_1^2) + s(s-1) \{E(u_1^2)\}^2 + 2s(s-1) \{E(u_1 v_1)\}^2. \quad (3)
 \end{aligned}$$

After some calculation we obtain

$$E(u_1^2 v_1^2) = a(c-a) \{ac(c-a) - c^2 + 3a(c-a)\} / c^4(c-1) \quad (4)$$

$$E(u_1^2) = (a/c)(1-a/c) \quad (5)$$

$$E(u_1 v_1) = -(a/c)(1-a/c)/(c-1). \quad (6)$$

Substitution of (4), (5) and (6) in (3), followed by substitution of (2) and (3) in (1), yields

$$E(X^2) = (c-1)^2 + 2(c-1)(1-1/s)$$

$$\text{or} \quad \text{var } X = 2(c-1)(1-1/s).$$

Hence the properties are established.

X is a quadratic form in the random variables $n_i - sa/c$ ($i=1, \dots, c-1$) which, as $s \rightarrow \infty$, become normally distributed with zero means and non-singular covariance matrix. With $E(X) = c-1$ this implies that X is asymptotically (with s) distributed as chi-squared with $c-1$ degrees of freedom, χ^2_{c-1} . Hence an approximate test for randomness at the 100 $\alpha\%$ level is provided by the critical region $X \geq X_\alpha$ where $\text{Prob}(\chi^2_{c-1} \geq X_\alpha) = \alpha$. The fact that $\text{var } X < \text{var } \chi^2_{c-1}$ for all s suggests that the exact level of significance will be somewhat higher or that the test will be conservative.

When $a=c-1$, the allocation leaves one cell unselected and

$$X = \sum (n_i - sa/c)^2 / (s/c) = \sum (m_i - s/c)^2 / (s/c)$$

where m_i is the number of times the i 'th cell is unselected. This is the ordinary expression for chi-squared when the units are unselected cells. More generally, X has the desirable property of being invariant under change of labelling of cells.

When $a=1$ or $c-1$, the approximation given by chi-squared is good provided $s \geq 5c$. It is likely that for these values of s the approximation will be satisfactory for intermediate values of a .

REFERENCE

- [1] KENDALL, M. G. (1952). *The Advanced Theory of Statistics*. I. Fifth edition. Griffin.

THE SALIENT VARIABLE SIMILARITY INDEX FOR FACTOR MATCHING

By RAYMOND B. CATTELL and ANDREW R. BAGGALEY
University of Illinois University of Wisconsin

The various techniques available for comparing factor-patterns may be conveniently classified according to the type of measurement scale for which each is most appropriate. If the factor loadings be regarded as based simply on a 'nominal' scale (i.e. on a dichotomous distinction between salient and non-salient loadings), then a quick working procedure can be derived which, being non-parametric, avoids the questionable assumption that the loadings are normally distributed. The test so obtained proves less powerful than a parametric test, but has special advantages both theoretical and practical.

I. THE NEED FOR A NEW TECHNIQUE

Earlier Proposals. In several areas of psychology in which factor analysis has been applied, we are entering on a new phase in which integration by the cross-matching of factors (usually reference vectors obtained in different investigations) becomes of paramount importance. A factor once found remains merely a hypothesis about a pattern; it is verified only after the pattern has been found again in well-defined circumstances. The architectonic growth of knowledge about factors and their meaning, which may now be confidently anticipated, can occur only to the extent that we develop new techniques for establishing significant factor invariance and for recognizing the modifications to be expected from one situation to another.

Surveys of the principal areas ripe for such integration were made in the personality area by Cattell [2] in 1946 and in the area of aptitude and achievement by French [9] in 1951. It is in connection with more recent surveys [7, 10] of the former field that the issue of precise matching has become crucial. The belief that factors may be identified from one study to another by mere inspection and impression has rightly been questioned by several experimenters, notably by Reyburn and Raath [14], who insist that "identity requires a more rigorous criterion than is commonly applied". The extreme form of this old belief—that, even when two studies have no variables in common, the factors can be matched by intuitions about their identity of meaning—must be written off as an endemic subjectivity-illusion of psychologists.

In an attempt to inaugurate the use of an objective index, Cattell [3] proposed in 1949 a specific and a general, a symmetrical and an unsymmetrical marker variable matching index. One of these was used to support the conclusion that substantial agreement existed among twelve factors found in

three successive factorizations. The present paper is a more specific development of that general approach in the light of closer reasoning and more extensive experience.

The Nature of the Problem. Six characteristics can be listed by which an oblique factor could be identified : (i) its loading profile on a given set of variables ; (ii) its correlation with other factors ; (iii) the amount of variance it extracts from a standard set of variables relative to other factors derived from those variables ; (iv) the company it keeps in a given factor analysis ; (v) the relation of the loadings to other sets found by different techniques (e.g. P- and T-technique) ; (vi) the way in which the factor pattern alters with experimental controls and sampling selection. The fourth characteristic implies that a factor cannot be given a particular interpretation if this has already been bestowed on another factor in the same study. The fifth and sixth criteria imply an examination of the first three in other settings. All these bases for interpretation deserve more explicit examination than they have hitherto received. Here we shall confine our development to the first alone, namely, the loading profile.

We assume that a factor analysis has been carried out with the same variables (and perhaps with others) on two random samples from the same population. For each sample the factors have been rotated to a position believed to correspond to a structure or resolution inherent in nature, by means of the principle of simple structure [16], criterion rotation [8], parallel proportional profiles [2], or some other technique. The question arises whether a certain factor in one sample can be considered the same as a factor in the other. An example illustrating these requirements—with 52 variables in common, 15 pairs of factors, and blind rotation for simple structure—has recently been examined by the methods to be suggested [5]. Strictly speaking, no factor derived from one sample can be mathematically the *same* as one derived from another. Our assumption is merely that a certain 'real' influence or dimension of variation, affecting several variables, is at work in both samples ; and we are asking whether the two patterns could be considered as due to this same influence operating in both cases.

Possible Techniques. To consider the matching of profiles in a basic fashion, it is necessary to ask (a) which of the usual scales—ratio, interval, ordinal or nominal ([15], pp. 23–30)—best applies to the loading measures, and (b) whether we are more concerned to get similarity of shape or of level.

As to scales, the assumptions in the techniques available suggest that the following are the most appropriate allocations.

<i>Scale</i>	<i>Technique</i>
Ratio	Parallel proportional profiles ;
Interval	Product-moment correlation ;
Ordinal	Tau ;
Nominal	Salient variable similarity index.

Besides parallel proportional profiles, another formula is available which treats factor loadings as falling on a ratio scale, namely Tucker's 'coefficient of congruence' [17], also proposed independently by Wrigley and Neuhaus under the name 'index of factorial similarity'. Regarding the second question, almost equal claims can be made for matching factors (i) on shape only and (ii) on the basis of absolute loading level attained. If the latter is adopted, the r_p coefficient [4] is essential. However, this, as well as the two ratio scale techniques, involves assumptions which space does not allow us to examine here. The method of parallel proportional profiles is considered in [6]; moreover, except where the two samples have been deliberately selected by different methods, it is less suitable to the matching problem than the indices we shall consider.

As to the four techniques listed in the table above, the first two are parametric devices : and these require the population distributions to be specified. The last two are strictly non-parametric. Tau [12] avoids the assumption of normality, which the product-moment correlation requires ; nevertheless, it seems inferior to our proposed index for two reasons. First, its calculation involves more work ; secondly, a probability value must be calculated in addition to tau, whereas the new index, s , is itself a probability value. The departure of the distribution of factor loadings from normality might be thought at first to affect the use of probability values calculated from r . It has, however, been shown [12, pp. 78-90] that such departures do not appreciably impair the variance ratio test. This finding permits the comparison of significance tests from r and s on empirical data.

The index discussed below was formerly called the 'marker similarity index', m [3] ; but a more appropriate name is the 'salient variable similarity index', s . 'Marker variables' are variables carried over, usually at the rate of two or three per factor, to mark the factor in a new study. To obtain the index proposed a certain number of variables will have to be selected out of those common to both studies. These 'salient' variables will be chosen according to the size of their factor loadings ; e.g. they may be the six, ten, etc., highest loadings. In certain cases it may turn out that no marker variables from the first study appear among the salient variables of the second. Precisely what proportion of the common variables should be considered salient is a difficult question. For a given pair of factors the index requires that the same number be taken for both. In theory it is not absolutely essential to keep this number constant for every pair ; but intuitively it seems wise to do so in order that the values found shall be comparable for all the pairs.

It has been suggested that the division between salient and non-salient variables be drawn at the point where the factor loadings cease to deviate significantly from zero. Strictly speaking, such a criterion is irrelevant in a non-parametric test. However, since nothing in the test itself specifies a particular point of division, such a criterion has an intuitive appeal. There is as yet no theoretical agreement as to the best way of testing significance ; but

the inclusion of a few random variables in the analysis will yield an empirical indication. Some of our colleagues have suggested that differences in the communalities should be taken into account. Our contention is, however, that the communalities are determined at the same time as the number of factors extracted : they apply to the study as a whole, and are determined as much by the non-common variables as by those common to the two studies. It would therefore seem that the communality problem is quite independent of the choice of a suitable proportion for the salient variables.

Measures of Factor Similarity. The earlier work carried to a certain stage the derivation of a measure for factor similarity from inspection of the salient variables. It consisted of two formulae [3, p. 137], rewritten below in a notation consistent with what is adopted here. The first gave the probability of at least c' of the g salient variables of the two factors being coincident as

$$s_{c'} = \frac{\binom{g}{c'} \binom{n-g}{g-c'} + \binom{g}{c'+1} \binom{n-g}{g-c'-1} + \dots + \binom{g}{g} \binom{n-g}{0}}{\binom{n}{g}}, \quad (1)$$

where n is defined as the total number of variables common to the two studies. The prime is used to distinguish the quantities in the earlier formulae from those in the present development. The second formula gave the probability of agreement of factors for the two studies as a whole as

$$p_{KN} = \sum_{z=N}^{z=K} w^z (1-w)^{K-z} \binom{K}{z}, \quad (2)$$

where K is the number of factors in each study, N is the largest number of exclusively matching pairs, w is $1 - (1 - s_{c'})^k$, and z is a summation index.

The argument for formula (1) is best perceived if we set aside consideration of the accumulation of probabilities for c' or more coincidences, and confine our attention to the probability of exactly x' coincidences. Formula (1) can be subdivided into several probabilities of the form

$$p_{x'} = \frac{\binom{g}{x'} \binom{n-g}{g-x'}}{\binom{n}{g}}. \quad (3)$$

We can then rewrite (1) as

$$s_{c'} = p_{c'} + p_{c'+1} + \dots + p_g = \sum_{x'=c'}^{x'=g} p_{x'}. \quad (4)$$

In developing the rationale let us consider both a particular numerical example and the general case which it illustrates. Consider the matching of two factors, each arising from the same 15 variables. The loadings are arranged in order of magnitude ; and from each list we designate as salient the variables

representing the highest six. This proportion is selected for the sake of exposition ; according to the criterion given above, the proportion would usually be smaller if the hyperplane were well determined. Having determined these six variables, we consider all the ways in which the 15 variables may arrange themselves with regard to the six highest loadings on the second factor. This gives the total number of possible events, which is the denominator of p_x . The number of possible combinations of 15 things taken six at a time is written $\binom{15}{6}$, or more generally $\binom{n}{g}$. The numerator represents the number of favourable events.

Suppose now that examination of the variables reveals that three of the six salients are the same in the two studies ; then $x' = 3$. Three of the highest six variables for the second factor must be drawn from the six highest on the first factor. The number of ways in which this can happen is $\binom{6}{3}$, or more generally $\binom{g}{x'}$. We must also account for the ways in which the other three of the highest six loadings on the second factor can be composed. There are nine non-salient variables available from the first factor. Therefore the number of possible combinations is $\binom{9}{3}$, or more generally $\binom{n-g}{g-x'}$.

II. THE NEW COEFFICIENT

Derivation. Formula (1) is theoretically correct if the g most highly loaded variables are selected *with regard to sign*. However, this forces us to choose either the g most highly positively loaded variables or the g most highly negatively loaded. In the cognitive area common sense often dictates an 'efficient' direction (usually that of the 'right answers'), and a positive manifold is thus secured. However in the areas of temperament and motivation the direction of scoring (e.g. whether the schizothyme or the cyclothyme gets the high score) is usually arbitrary. We developed therefore a new formula for factor loadings ordered *without regard to sign*. After selecting the salient variables, the signs are again taken into account in selecting coincident variables. The salient variables will now be those with the g highest absolute values of factor loading. In general for each factor these will not be exactly the same as the g variables used in formula (1). We compare the lists for the two factors, and pick out the c variables which are common to the two lists and have loadings *with the same sign*. In general this number will differ from c' used in formula (1). Therefore the prime has been dropped. Corresponding to formula (3) we have the new formula¹,

¹ We are indebted to C. R. Blyth, of the Department of Mathematics, University of Illinois, for developing this part of the formula and for deriving a simpler expression than that which we had originally reached.

$$p_x = \frac{\binom{g}{x} \sum_{y=0}^{y=g-x} \left[\binom{n-g}{g-x-y} 2^{g-x-y} \binom{g-x}{y} \right]}{2^g \binom{n}{g}}, \quad (5)$$

where y is a summation index.

In illustrating formula (5) we shall keep $n = 15$ and $g = 6$, but shall suppose that only two salient variables match with the same sign in the two factors, i.e. $x = 2$. In the denominator 2^g is necessary because the g salient variables from the first factor may appear in the second list with *either* positive or negative signs. For the numerator two of the highest six variables for the second factor must be drawn from the six highest for the first, thus accounting for $\binom{6}{2}$, that is, $\binom{g}{x}$. In this case they must have the *same* sign.

In regard to the other four of the highest six loadings on the second factor the situation has become more complex. As before they can represent the nine non-salient variables from the first factor, and can appear in the highest six of the second factor with *either* sign. But now they can also represent variables which were in the highest six of the first factor, and which appear in the highest six of the second factor with *opposite* sign. These two classes of variables can appear in several proportions, 4 : 0, 3 : 1, 2 : 2, 1 : 3, or 0 : 4. The symbol y in formula (5) indicates the number of the four loadings which represent variables of the second class. The term $\binom{n-g}{g-x-y}$ accounts for the number of ways of distributing the first class of variables, those not in the highest six of the first factor, if no sign changes occurred; and the term 2^{g-x-y} accounts for the possible sign changes in the first class of variables. The term $\binom{g-x}{y}$ accounts for the number of ways of distributing the second class of variables, those which are in the highest six of the first factor. With oblique factors the situation will become more complicated, and the development of an appropriate formula will doubtless require further advances in multivariate analysis.

Computation Procedure. As often happens, the formula which is simplest algebraically is not the simplest arithmetically, either in terms of labor or of understanding by a computing clerk. Accordingly, when n exceeds 10, the formulae now to be developed are recommended in preference to equations (4) and (5). In the example we have taken, as with most empirical data, c is less than half of g . Hence computation will be simplified if the values for p_x are summed from $x = 0$ to $x = c - 1$ and subtracted from unity. Formula (4) can then be rewritten as

$$s_c = 1 - \sum_{x=0}^{x=c-1} p_x. \quad (6)$$

The primes have been dropped because we are now taking signs into consideration. We can simplify formula (5) by using logarithms of factorials. First, all the factors are brought into the summation, so that

$$p_x = \sum_{y=0}^{y=g-x} \frac{2^{g-x-y} \binom{g}{x} \binom{n-g}{g-x-y} \binom{g-x}{y}}{2^g \binom{n}{g}}. \quad (7)$$

To evaluate (7) in terms of factorials, we have the general formula,

$$\binom{a}{\beta} = \frac{a!}{(a-\beta)! \beta!}.$$

Taking the logarithm of each side, we have $\log \binom{a}{\beta} = \log a! - \log \beta! - \log (a-\beta)!$. We then obtain from (7)

$$\begin{aligned} & \log \left[\frac{2^{g-x-y} \binom{g}{x} \binom{n-g}{g-x-y} \binom{g-x}{y}}{2^g \binom{n}{g}} \right] \\ &= [(g-x-y) \log 2 - g \log 2] \\ &+ [\log g! - \log x! - \log (g-x)!] \\ &+ [\log (n-g)! - \log (g-x-y)! - \log (n-2g+x+y)!] \\ &+ [\log (g-x)! - \log y! - \log (g-x-y)!] \\ &- [\log n! - \log g! - \log (n-g)!]. \end{aligned} \quad (8)$$

Now we combine similar terms and separate those involving y , thus obtaining

$$\begin{aligned} & \log \left[\frac{2^{g-x-y} \binom{g}{x} \binom{n-g}{g-x-y} \binom{g-x}{y}}{2^g \binom{n}{g}} \right] \\ &= [2 \log g! + 2 \log (n-g)! - \log n! - \log x!] \\ &- [(x+y) \log 2 + 2 \log (g-x-y)! + \log (n-2g+x+y)! + \log y!]. \end{aligned} \quad (9)$$

Equation (9) can be written in a less complex fashion by defining three new symbols :

$${}_x a = 2 \log g! + 2 \log (n-g)! - \log n! - \log x!, \quad (10)$$

$${}_x b_y = (x+y)(0.301) + 2 \log (g+x-y)! + \log (n-2g+x+y)! + \log y!, \quad (11)$$

and

$${}_x t_y = {}_x a - {}_x b_y. \quad (11)$$

Now for all t 's in which the sum of x and y is the same, their corresponding terms have identical values except for $\log x!$ and $\log y!$, which are simply interchanged with constant sign. Therefore ${}_ut_v = {}_xt_u$. It follows that (9) can be written

$${}_yt_x = \log \left[\frac{2^{g-x-y} \binom{g}{x} \binom{n-g}{g-x-y} \binom{g-x}{y}}{2^g \binom{n}{g}} \right]. \quad (13)$$

Taking the antilogarithm of both sides of (13), we have

$$\text{antilog } {}_xt_y = \left[\frac{2^{g-x-y} \binom{g}{x} \binom{n-g}{g-x-y} \binom{g-x}{y}}{2^g \binom{n}{g}} \right]. \quad (14)$$

Combining (7) and (14) give the formula,

$$p_x = \sum_{y=0}^{y=g-x} \text{antilog } {}_xt_y. \quad (15)$$

A Numerical Example. Formulae (10), (11), (12), (15), and (6) enable us to calculate the probability for any particular example. We shall now show the use of these formulae in the example mentioned above. A table of logarithms of factorials is provided by Arkin and Colton [1, p. 109]

$$n=15, g=6, n-g=9, n-2g=3$$

$x=0, {}_0a=4.718$			
y	${}_0b_y$	${}_0t_y$	antilog ${}_0t_y$
0	6.492	8.226-10	0.017
1	5.839	8.879-10	0.076
2	5.742	8.976-10	0.095
3	6.094	8.624-10	0.042
4	6.888	7.830-10	0.007
5	8.190	6.528-10	0.000
6	10.223	4.495-10	0.000

$x=1, {}_1a=4.718$			
y	${}_1b_y$	${}_1t_y$	antilog ${}_1t_y$
1	5.441	9.277-10	0.189
2	5.617	9.101-10	0.126
3	6.286	8.432-10	0.027
4	7.491	7.227-10	0.002
5	9.445	5.273-10	0.000

$x=2, {}_2a=4.417$			
y	${}_2b_y$	${}_2t_y$	antilog ${}_2t_y$
2	5.809	8.608-10	0.041
3	6.889	7.528-10	0.003
4	8.746	5.671-10	0.000

$x=3, {}_3a=3.940$			
y	${}_3b_y$	${}_3t_y$	antilog ${}_3t_y$
3	8.144	5.796-10	0.000

antilog ${}_xt_y$							
y	x						
	0	1	2	3	4	5	6
0	0.017	0.076	0.095	0.042	0.007	0.000	0.000
1	0.076	0.189	0.126	0.027	0.002	0.000	
2	0.095	0.126	0.041	0.003	0.000		
3	0.042	0.027	0.003	0.000			
4	0.007	0.002	0.000				
5	0.000	0.000					
6	0.000						
p_x	0.237	0.420	0.265	0.072	0.009	0.000	
${}_cp_x$	0.237	0.657	0.922	0.994	1.003	1.003	
s_c	0.76	0.34	0.08	0.01	0.00	0.00	
c	1	2	3	4	5	6	

Thus the probability of getting two or more matches by chance from six salient variables out of fifteen common variables is 0.34. To illustrate the procedure, sample values of ${}_xa$, ${}_xb_y$, and ${}_xt_y$ are calculated below :

$${}_1a = 2 \log 6! + 2 \log (15-6)! - \log 15! - \log 1! = 4.718.$$

$${}_1b_3 = (1+3)(0.301) + 2 \log (6-1-3)! + \log [15-2(6)+1+3]! + \log 3! = 6.286.$$

$${}_1t_3 = 4.718 - 6.286 = 8.432 - 10.$$

The p_x values are column sums. They are accumulated across in the cp_x row of the last table. The sums are then subtracted from unity to obtain the s_c values, since the probability of getting zero or more matches is 1.00. Note that the triangular matrix is symmetrical about the principal diagonal, because antilog ${}_ut_v$ equals antilog ${}_vt_u$. This means that when a column of entries has been calculated, the corresponding row can be written down immediately. Also the calculations for a column can safely be stopped when the row is reached for which ${}_xt_y$ is less than 6.699-10 (the logarithm of 0.0005). This also means a saving in columns ; in the above example, for instance, the columns for $x=5$ and $x=6$ are unnecessary.

The above example has been worked out for all values of c when $g=6$. In actual practice the computer can sometimes stop computing p_x values even sooner, according to the highest empirical value of c for a given pair of studies. In the above example, if no more than three matches were found for any pair of factors, then the columns for $x=3$ and $x=4$ would also be unnecessary. (Note that c always exceeds x by one for a column of the table.)

A few more comments may help the computer. Because all values of $\text{antilog } {}_x t_y$ are probabilities, they must all be less than unity. Therefore all values of ${}_x t_y$ must be negative; and all values of ${}_x b_y$ must be greater than the corresponding ${}_x a$. If the computer obtains a ${}_x b_y$ less than the corresponding ${}_x a$, he should stop, and check his calculations. The complementary form of the negative ${}_x t_y$, obtained from the calculator dials is retained, since it facilitates finding the antilogarithm.

Since three parameters, n , g , and c , are involved in the calculation of s_c , a complete table would be three-dimensional. To present a single two-dimensional table, we have taken g as a constant fraction of n , namely, one-sixth—the value used in the empirical example of the next section (cf. Table I). However, the use of s_c is not restricted to any one fraction of this kind. If an investigator decided to use some other constant fraction, he could make another such table. If he decided to vary the fraction between factor pairs, he could use the formulae directly. Table I gives the values of s_c for all possible values of c when $g < 11$ and $n = 6g$. The 0.00 entries are not exactly zero; they are values less than 0.005.

III. BEHAVIOUR OF THE SIMILARITY INDEX IN ACTUAL EXAMPLES

A recent survey (7) applying the above and other matching indices to a variety of factors in the personality field permits us to make some empirical observations on its practical usefulness and its consistency in relation to other evidence. In all, this survey examined about 150 individual factors reported as patterns from single studies, and concluded that most of them are varying manifestations of 19 different psychological influences which underlie them. 104 of the factors were arranged in 19 matrices, each exhibiting all possible interrelations of three to seven phenomenal factors matched from different researches. For each pair of factors the similarity of factor pattern (reference vector structure was actually used) was tested (i) by the product-moment r and (ii) by the s coefficient. For any two studies the number of common variables (n) ranged from 15 to 100. The selection of salient variables was made by taking g always as one-sixth (rounded off to an integer) of n : i.e. for each factor we considered the highest sixth (as judged by the absolute value of factor loading) of the variables shared with the other study.

The s coefficient provides a one-tailed significance test. Ordinarily product-moment r is tested by a two-tailed test; here, however, since we had to specify the direction of scoring the factors in calculating s , the corresponding r must also be tested by a one-tailed test. For this purpose we used the t

function given by Johnson [11, p. 51]; and, as explained above, we assumed that the departure of the distribution of the loadings from normality would not seriously affect its use.

To compare the parametric and non-parametric data for each s , we looked up the base line value which cut off that proportion of the normal curve above it. For example, the normal deviate for $s=0.20$ is $+0.84$; for $s=0.60$ it is -0.25 . Table II shows the scatter of t from r against the normal deviate from s . The correlation between two variables from this grouped data is $+0.54$. Inspection of the table shows [3, p. 139] that r is the more sensitive measure. There were some cases in which r was high, even when there were few (or even no) matches of the salient variables. Inspection of the paired loadings showed that variables with low loadings (non-salient variables) sometimes still matched in sign. On the other hand, nearly all cases in which s indicated a good match were substantiated by r . Thus, in mathematical terms, t_r proves to be a more 'powerful' test than s .

Since we purposely selected the best matches, the r and s values given in Table II are not directly relevant to the question of the probabilities involved when a factor from one study is matched in turn with *all* the factors from another study. Strictly speaking, the value of s (like p for the t calculated for r) represents the probability of matching for a pair of factors *specified in advance*. The answer to the further question therefore would seem to lie in an approach similar to that used in deriving equation (2).

In general, a parametric approach (or perhaps *tau*) seems advisable when a few isolated pairs of columns of factor loadings are under consideration. However, s saves labour when several pairs are being matched between two or more factor analyses. Here there will often be several pairs of factors from a given pair of analyses which will give the same number of matches, and therefore the same s value. A parametric approach would require a separate correlation (or similar index) for each. Also, of course, the non-parametric approach avoids assumptions concerning the form in which loadings are distributed.

Summary. 1. The matching of factors depends on more than their patterns, but in this paper we have considered only the patterns. Various techniques of pattern comparison, corresponding to Stevens' scales of measurement, have been considered. Theoretical expressions and computational formulae are derived for a non-parametric approach, in which the factor loadings are treated as falling on a nominal scale (based on a dichotomous classification into salient and non-salient loadings).

2. The resulting s coefficient and the product-moment correlation technique are compared on empirical data. An agreement corresponding to a correlation of $+0.54$ is reported. The s test proves less powerful than a parametric test, but avoids the questionable assumption of normality.

3. The salient variable similarity index, s , is recommended for evaluating several matches from two or more large studies.

[illegible]

TABLE II. SCATTER DIAGRAM FOR FACTOR MATCHING INDICES

z_s	t_r													Total
	-2.50	-2.00	-1.50	-1.00	-0.50	0.00	+0.49	+0.50	+0.51	+1.00	+1.50	+2.00	+2.50	
+2.51 to + ∞						1					1	1	4	+ ∞
+2.01 to +2.50				1				4		1	6	8	2	+2.51
+1.51 to +2.00								1	1	1	3	3	6	14
+1.01 to +1.50								1	1	3	4	4	1	7
+0.51 to +1.00						3		5	3	2	6	5	7	18
+0.01 to +0.51			1	1		5		9	2	1	1	1	1	12
-0.49 to 0.00					1	1		1	1	1	1	1	1	8
-0.99 to -0.50						1								6
-1.49 to -1.00														13
-1.99 to -1.50														2
-2.49 to -2.00														
- ∞ to -2.50	1	1	2	3	3	13	20	1	1	5	5	3	2	60
Total	1	1	3	5	4	24	40	9	9	27	27	25	24	240

REFERENCES

- [1] ARKIN, H., and COLTON, R. R. (1950). *Tables for Statisticians*. New York : Barnes and Noble (College Outline Series).
- [2] CATTELL, R. B. (1946). *The Description and Measurement of Personality*. New York : World Book Company.
- [3] CATTELL, R. B. (1949). A note of factor invariance and the identification of factors. *Brit. J. Psychol., Statist. Sec.*, II, 134-139.
- [4] CATTELL, R. B. (1949). r_p and other coefficients of pattern similarity. *Psychometrika*, XIV, 279-298.
- [5] CATTELL, R. B. (1953). *Personality Structure in Objective Tests : A Preliminary to the Development of Psychiatric Screening*. San Antonio : AF School of Aviation Medicine.
- [6] CATTELL, R. B. (1957). *Personality and Motivation : Structure and Measurement*. New York : World Book Co.
- [7] CATTELL, R. B. (1955). Factor rotation for proportional profiles : analytical solution and an example. *Brit. J. Psychol., Statist. Sec. VIII*, 83-92.
- [8] EYSENCK, H. J. (1950). Criterion analysis : an application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.*, LVII, 38-53.
- [9] FRENCH, J. W. (1951). *The Description of Aptitude and Achievement Tests in Terms of Rotated Factors*. Chicago : Univ. Chicago Press.
- [10] FRENCH, J. W. (1953). *The Description of Personality Measurement in Terms of Rotated Factors*. Princeton : Educational Testing Service.
- [11] JOHNSON, P. O. (1949). *Statistical Methods in Research*. New York : Prentice-Hall.
- [12] KENDALL, M. G. (1948). *Rank Correlation Methods*. London : C. Griffin.
- [13] LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education*. Boston : Houghton Mifflin.
- [14] REYBURN, H. A., and RAATH, M. J. (1949). Simple structure : a critical examination. *Brit. J. Psychol., Statist. Sec.*, II, 125-133.
- [15] STEVENS, S. S. (1951). Mathematics, measurement, and psychophysics. Ch. 1 in S. S. STEVENS (Ed.), *Handbook of Experimental Psychology*. New York : Wiley.
- [16] THURSTONE, L. L. (1947). *Multiple-Factor Analysis*. Chicago : Univ. Chicago Press.
- [17] TUCKER, L. R. (1951). A method for synthesis of factor analysis studies. *PRS Report No. 984, Personnel Research Section, Dept. of the Army, Washington, D.C.*

TAYLOR ON 'LAW' AND 'THEORY' IN PSYCHOLOGY

By G. W. PILKINGTON

Department of Psychology, University of Sheffield

This paper criticizes the ideal of 'law' and 'theory' in psychology put forward by Taylor. It is suggested that his first two requirements for 'genuine' scientific laws are so stringent as to be unrealistic, and that the second two, though sound in principle, are too vague for it to be clear what particular psychological laws they are intended to exclude. Finally, reasons are adduced for rejecting what Peters has called 'dogmatic methodism', of which Taylor's proposals are typical examples.

I. TAYLOR'S FOUR REQUIREMENTS

Mr. J. G. Taylor's recent papers [15, 16] on the nature of scientific method in psychology are forceful and interesting, but at the same time extremely provocative, as recent contributions to this *Journal* [3, 4, 13, 19] sufficiently show. In this note I wish to discuss a central contention of his papers—his notion of what constitutes a genuine scientific theory or law in psychology: for from this flow most of his other contentions,¹ including his derogatory view of statistical psychology and experimental design.

Taylor lays it down that in psychology 'genuine' or 'proper' scientific theories or laws² should have the following properties. (i) They should be highly general and abstract, that is, they should be much more than mere generalizations of observed relations between phenomena. (ii) They should be universal in the sense that they admit of no exceptions and are not mere statements of probability; thus he claims that "a law of nature can never be stated in the form of a statistical law". (iii) They should be consistent with the laws and empirical findings of other sciences. (iv) They should be logically consistent, that is, no part of the theory or law should contradict any other.

Here is a formidable set of requirements: for, according to Taylor, they entail that "there is no body of doctrine to which every psychologist can subscribe as there is in other scientific disciplines". "The one thing an applied psychologist cannot and dare not apply," we are told, "is *psychology* in any forms in which it is at present understood." But if the consequences of these requirements are so radical, then surely their credentials should be more carefully examined.

¹ Though not his distinction between 'psychology' and 'behavioural science' which is extremely puzzling: I am sure that many psychologists would welcome a more detailed elucidation.

² I do not differentiate between 'theories' and 'laws', not because I believe this cannot, or should not, be done, but because a distinction between them is not vital to the purpose of this paper.

II. PHYSICS AS THE PARAGON OF THE SCIENCES

Taylor's first requirement would seem to owe much to what has come to be called the 'hypothetico-deductive' view of science. This is clearly implicit in his statement that "a valid theory . . . must be so formulated that, when non-contradictory values are assigned to the several variables and parameters of the system, it is possible to draw specific deductions that can be tested by observation and experiment". He emphasizes that scientific laws and theories involve much more than the mere "counting of heads". A law is no mere empirical generalization, like "all stockbrokers wear pin-striped trousers", but exhibits a high degree of abstractness. Taylor makes this plain by his examples from Newton. In his view, a scientific law is a conceptual device which enables us to deduce, from a conjunction of (i) statements specifying the initial conditions and (ii) the rules of inference formulated by the law, a whole series of further properties which the phenomena envisaged by the law should logically possess. His methodological ideal comes very close to that expounded by Hull in the well-known first chapter of his *Principles of Behaviour* [9]. Now if phrases like 'scientific law' and 'scientific theory' are to be used so as to apply *only* to complicated conceptual structures of this kind, then Taylor obviously is right in holding that there are no 'proper' laws or theories in psychology. But is it reasonable thus to restrict the use of these terms? The working scientist does not do so. What then is to be gained by advising him to change his linguistic habits?

My own reply would be—little or nothing. To begin with, few branches of science have laws and theories of this highly abstract and formalized kind. Geography, geology, botany, zoology, physiology, anatomy, are recognized sciences. Yet none has developed, or even strives to develop, classical hypothetico-deductive theories. To use language as Taylor recommends would mean that certain parts of physics, chemistry, and astronomy alone 'really' possess 'proper' scientific laws and theories. We should then be reduced to the sad position of Hull, lamenting that the only examples he *can* give of 'proper' scientific explanation must be drawn from the physical sciences [9, p. 3]. But, as has often been remarked, psychologists tend to be over-influenced by the prestige of the pattern of explanation achieved by the exact physical sciences, and have paid too little attention to the logical features of explanation in the biological and social sciences. And surely psychology resembles these sciences much more than it resembles physics. Taylor is the latest of many psychologists to try to fashion psychology in the false image of physics.

Another objection to Taylor's requirement is the wide disparity between psychological phenomena and those studied by physicists. Psychological phenomena are highly complex in a number of ways: (a) it is difficult even to describe them objectively (witness the attempts of many behaviour theorists to develop a 'data language' for psychology); (b) it is harder still to quantify

them ; (c) they are the product of numerous heterogeneous variables ; and (d) the variables themselves are difficult to control experimentally. The phenomena of physics, on the other hand, are very different. Thus, it is far easier to manipulate them experimentally and formalize them mathematically. Taylor invokes the analogy of the solar system to support his conception of scientific law [pp. 109–110]. But his argument obscures important differences between solar systems and human behaviour and experience. Solar systems are relatively closed systems. The external variables impinging on them are few and relatively constant. And it is just this property of solar systems that is responsible for the fact that we have been able to develop abstract laws to explain and predict their behaviour with such high efficiency. Human beings do not form closed systems. Their behaviour is the result of a vast number and variety of factors, both internal and external. Moreover, human beings do what no solar system does : they store within them the effects produced by certain variables acting on them in the past. Hence in psychology the task of establishing abstract laws of great generality is far more difficult than in astronomy or classical physics—perhaps even impossible, as has been most plausibly argued by M. M. Braithwaite [1]. Of course, Taylor has a perfect right to base his researches upon the presumption that highly abstract laws in psychology *can* sooner or later be obtained. But he cannot insist that we all do the same.

III. THE REJECTION OF STATISTICAL LAWS

Taylor's second requirement—that a scientific law should admit of no exceptions—similarly indicates the extent to which the achievements of physics have influenced his paradigm of explanation. Here once more we see the assimilation of all scientific laws and theories to the particular pattern that happens to have developed in Newtonian physics. But again we may ask how far such a requirement is reasonable? After all, not even physics nowadays employs deterministic laws alone. As previous papers in this journal have pointed out, the theory of quantum mechanics, which explains the phenomena of atomic physics, is itself a statistical theory incompatible with a purely deterministic explanation of the phenomena involved. That being so, it is surely a little paradoxical to suggest that statistical theories are not 'genuine' theories at all. Moreover, in meteorology, and in the social, biological, and medical sciences, many laws and theories are statistical or probabilistic in form. No doubt, it may be argued that all statistical laws will eventually be superseded by laws that are fully deterministic. But what would count as 'superseding'? Perhaps all that is meant is that statistical laws will some day be *explained* by deterministic laws ; for example, the statistical law relating cigarette-smoking and lung cancer may eventually be explained in terms of physiological or biochemical factors. Nevertheless, the statistical law would still remain useful for practical purposes : thus, we could cite it when advising against cigarette-

curiosity is a statistical rather than deterministic law. Obvious examples occur in sociology and economics : e.g., findings about the leisure activities or the expenditure of people belonging to different social classes. Similar considerations apply to psychology. The laws of learning and personality development may some day be sufficiently well advanced to explain the different details of reaction of a hundred people, who, however, respond in broadly similar ways to a given situation. Such a situation might obtain, say, in some perception experiment ; yet here it is perfectly possible for psychologists to be interested in *general tendencies*. This is just as scientifically legitimate as a predominant interest, such as Taylor seems to show, in explaining *individual differences* in perception in objectively the same situation. There is no reason why such tendency statements should not be dignified by the name of 'law' and in fact they often are. Do we not commonly speak of the Gestalt 'laws' of perception?

The whole notion then that relations between phenomena expressed in statistical terms are somehow at most pragmatically useful and in *no* circumstances worthy of the name of 'laws' must be rejected. They occupy a central place in many of the biological and social sciences—and even in physics. Moreover, the analysis of statistical laws and hypotheses is an important preoccupation of many contemporary philosophers of science, for example, Braithwaite [2]. None of this, of course, is to argue "that psychology is by its very nature an inexact science", as some of those whom Taylor castigates apparently do. In my view, it is pointless dogmatism to say that psychology (or any other field of enquiry for that matter) is *by its very nature* probabilistic, or deterministic, or teleological, or anything else whatever. If someone believes, as Taylor does, that ultimately all the phenomena which interest psychologists will be explainable in terms of deterministic laws, the most he can do is play his hunch. This simply means that he organizes his research so as to try to establish deterministic rather than statistical laws. What, however, he cannot do is lay down *a priori* the form psychological laws ultimately *must* take. No more can his rival who favours probabilistic laws. No one, without evidence, can say what the world *must* be like ; yet that is precisely what Taylor does. Commenting on the view that psychology is by its very nature an inexact science, he says "this seems to imply a belief, not merely that nobody has yet found the laws that will account *exactly* for all types of results, but that no such laws exist in nature. *But this is absurd!* Mind constitutes part of the system of nature ; and *it must be lawful*" (my italics). Yet it is certainly not self-contradictory to say that *some* of the phenomena of psychology are interrelated in terms of proportions rather than of uniformities ; consequently there is no reason for dismissing the supposition as intrinsically absurd. Whether or not mind is lawful, and if so, in what ways and to what extent, is a question of empirical fact, not of *a priori* assertion. No doubt the lawfulness, in some sense or other, of psychological phenomena is an implicit working presumption adopted by all empirical psychologists. smoking. Furthermore, there are occasions when what satisfies our intellectual

But this is quite different from declaring that there *must* be absolute lawfulness, especially lawfulness of a particular sort.

IV. INTERNAL AND EXTERNAL CONSISTENCY

Taylor's last two requirements present less difficulty. Of course theories and laws in psychology must be consistent with the laws and empirical findings of other sciences, and of course they must be logically consistent. We begin to feel doubts about the importance of the first requirement only because Taylor provides no convincing examples of extant theories and laws which contravene it. He assures us that psychologists "spin any kind of theory . . . without stopping to consider whether it is consistent with neurology or with behaviour theory", that they discover a "wide gulf" separating *Geistwissenschaft* from *Naturwissenschaft*, and that they have developed bad habits which prevent them from "seeing mind as an integral part of an organic system having the property of ultrastability". But these vague charges are simply not enough. What is wanted are specific examples. Taylor's remarks, as they stand, serve merely to confuse us still further about the distinction he draws between psychology and behavioural science. Most of us think of behaviour theory as *part of* psychology, not as some other discipline with which psychology must be consistent. Later Taylor asserts that "there can be no science of mind apart from the science of behaviour". But, in view of his earlier sharp distinction between the two, what precisely can this mean?

As to the requirement that laws and theories should be self-consistent, we are told that "in the science of behaviour it is customary to make some effort to satisfy this requirement; but in psychology it is hard to find a theory that does not violate it". Again there is this puzzling distinction between psychology and the science of behaviour, and again a dearth of specific examples. One is apparently provided, namely, that of the nativist-empiricist quarrel. But here it is hard to know *what* Taylor is trying to say. The straightforward interpretation is that he is assimilating 'innate' and 'acquired' to the status of logical contradictions: just as we cannot say "the pencil is both black and white all over", so also we cannot say "the ability to recognize the pitch of a musical note is both innate and acquired". But there is at least one legitimate sense in which we *can* say this: when we wish to put forward some such hypothesis as that "the ability to identify pitch depends partly on factors such as possessing a normal cochlea, auditory nerve, auditory cortex, and so on (in general, on inborn determinants), and partly on having been exposed to a musical environment in which fine pitch discrimination was repeatedly reinforced (in general, on acquired determinants)". Possibly Taylor is suggesting that a mutually exclusive usage of 'innate' and 'acquired' *ought* to be adopted because of certain factual considerations. But if so, then we have a right to expect from him at least some brief outline of what these factual considerations are. And surely we also deserve some indication that it is *he* who is recommending a change in

linguistic practice, instead of the stern, schoolmasterly tone of accusing us all of some logical howler.

V. CONCLUSION

Finally, I should like to develop a view implicit throughout this paper, namely, that it is a mistake to try to lay down in advance the form that *all* explanation in psychology (or, for that matter, in any science) *must* ultimately take. Peters [11, pp. 716f.] has drawn attention to a practice which he aptly describes as "dogmatic methodism". By this he means legislating on allegedly methodological grounds about the sort of course psychology *ought* to pursue if it is to become 'really' or 'truly' scientific, and so make effective progress. Examples are provided in abundance by the classical schools of psychology in the '20's and '30's, and by their slightly less vociferous successors in more recent times. Taylor is merely our latest dogmatic methodist: for he enthusiastically instructs us about the form laws and theories in psychology ought to take in precisely the way that Peters deprecates. But surely Peters is right in suggesting that psychologists have already suffered a surfeit from such "methodological recipes".

In my previous contribution [13] I drew attention to various forms that dogmatic methodism may take, and doubtless others could be distinguished. All are equally objectionable. For human behaviour and experience cannot, for ever and always, *best* be explained in any *one* particular kind of way. If it were possible to offer more than a few hopelessly vague remarks about 'explanation', it would be to say that an important aspect of it involves establishing 'if-then' relations between phenomena of an invariant or probabilistic kind. But the phenomena of psychology are related to a multitude of different factors, ranging from the biochemical and neurological to the social and personal. Once established, such relations can themselves be related to each other in different ways or to laws concerned with different phenomena. Thus, a number of different explanatory systems, i.e. networks of interrelated 'if-then' relations, are possible in psychology. The same is true of other sciences concerned with complex, molar phenomena. The only exception arises when we investigate some relatively closed and homogeneous system. In general, however, it is a mistake to think that there can be a single 'true', 'real' or 'proper' explanation of a phenomenon or set of phenomena. Out of the array of conditions under which a phenomenon regularly occurs various principles may be selected to explain it; which is most appropriate must depend on our interests and purposes at the moment. The ideal of explanation that has worked well for one field of enquiry, physics, for example, may be quite unsuited for another, for example, psychology or sociology. More recently, Peters [12] has argued that all-inclusive, highly general theories of the kind Taylor apparently wants, and which physics has achieved in the work of Newton and Einstein, are logically impossible in psychology. Indeed, after skilfully exhibiting the wide variety of different kinds of explanation possible for human

traits and actions, Peters concludes that "the role of the Galileo of psychology must be for ever unoccupied" (p. 156).

To claim, then, that one type of explanation in psychology is the only 'genuine' kind is at best a rationalization of our own research preferences, or of the model of explanation we happen to be working with in our own thinking; at worst, it is mere pseudo-methodology. This is not to say that *nothing* valuable can be, or has been, said about the conditions a scientific explanation should satisfy: there is, for example, the requirement of intersubjective confirmability for scientific statements, as suggested by Popper [14]. Other conditions are discussed by Braithwaite [2], Burt [3], Hospers [8], Toulmin [17] and Gallie [7]. The latter's extremely instructive paper should be compulsory reading for all psychologists prone to lay down the law about the 'true' nature of science. The philosophical and logical analysis of explanations (laws, theories, etc.) *as they are actually used in the various sciences* may on occasion prove useful to the working scientist: it may help him to solve those problems that depend on conceptual issues. But this is very different from dogmatic methodism, which is mere bogus philosophy of science. Such restrictive legislation cannot benefit the psychologist one whit. It simply distracts him from his primary task—thinking up and testing fresh and original hypotheses about the determinants of human behaviour and experience. Certainly, as Taylor says, no amount of immaculate, but niggling, experimentation will make up for lack of such ideas; but neither will a blind attempt to imitate the procedures of Newton and Einstein.

REFERENCES

- [1] BRAITHWAITE, M. M. (1949). Causal laws in psychology: II. *Proc. Arist. Soc.* Supp., Vol. XXIII, 45–60.
- [2] BRAITHWAITE, R. B. (1953). *Scientific Explanation*. Cambridge: University Press.
- [3] BURT, C. (1958). Definition and scientific method in psychology. *Brit. J. Statist. Psychol.*, XI, 31–69.
- [4] BURT, C. and GREGORY, W. L. (1958). Scientific method in psychology: II. *Brit. J. Statist. Psychol.* XI, 105–128.
- [5] DREVER, J. (1952). *A Dictionary of Psychology*. Harmondsworth: Penguin.
- [6] ENGLISH, H. B. and ENGLISH, A. C. (1958). *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms*. New York: Longmans, Green & Co.
- [7] GALLIE, W. B. (1957). What makes a subject scientific? *Brit. J. Philos. Sci.*, VIII, 118–139.
- [8] HOSPERS, J. (1946). On explanation. *J. Philos.*, 43, 337–356.
- [9] HULL, C. L. (1943). *Principles of Behaviour*. New York: Appleton-Century Crofts.
- [10] MCKELLAR, P. (1957). *Imagination and Thinking*. London: Cohen & West.
- [11] PETERS, R. S. (1953). *Brett's History of Psychology*. London: Allen & Unwin.
- [12] PETERS, R. S. (1958). *The Concept of Motivation*. London: Routledge & Kegan Paul.
- [13] PILKINGTON, G. W. (1958). Scientific method in psychology: III. *Brit. J. Statist. Psychol.*, XI, 129–132.

- [14] POPPER, K. R. (1959). *The Logic of Scientific Discovery*. London : Hutchinson.
- [15] TAYLOR, J. G. (1958). Experimental design : a cloak for intellectual sterility. *Brit. J. Psychol.*, 49, 106-116.
- [16] TAYLOR, J. G. (1958). Scientific method in psychology : IV. *Brit. J. Statist. Psychol.*, XI, 133-136.
- [17] TOULMIN, S. (1953). *The Philosophy of Science*. London : Hutchinson.
- [18] WARREN, H. C. (1934). *Dictionary of Psychology*. Cambridge : Houghton Mifflin.
- [19] WILLIS, T. R. (1958). Scientific method in psychology : I. *Brit. J. Statist. Psychol.*, XI, 97-104.

LOGICAL POSITIVISM AND THE CONCEPT OF CONSCIOUSNESS

Critical Notice

By CYRIL BURT

University College, London

Logical Positivism. Edited by A. J. Ayer. Glencoe (Illinois) : The Free Press ; London : George Allen & Unwin. 1959. Pp. viii+455. 48s.

The Vienna Circle. This volume is the second in a new series, edited by Professor Paul Edwards of New York and called the 'Library of Philosophical Movements'. The name which gives the book its title was coined primarily to describe the standpoint of a group of Austrian thinkers who called themselves *Der Wiener Kreis*. To begin with they formed a club rather than a movement, and their meetings appear to have started shortly after World War I when Moritz Schlick arrived from Kiel to take up the chair of philosophy at Vienna. Schlick had already published his *Allgemeine Erkenntnislehre*, and in 1921 another Austrian, Ludwig Wittgenstein, brought out his *Logisch-Philosophische Abhandlungen*—both of which strongly influenced the views of the group. But their real progenitors were the nineteenth century empiricists, such as Comte, Mill, and their British followers, and the two great Viennese physicists, Mach and Ludwig Boltzmann.

Besides Schlick the chief members of the Circle included representatives of philosophy and logic like Rudolf Carnap, Otto Neurath, Herbert Feigl, Friedrich Waismann, of mathematics like Kurt Gödel, and of natural science like Philipp Frank. In 1929 they published their formal manifesto—*Der Wiener Kreis : Wissenschaftliche Weltauffassung* ; and in the following year they acquired a periodical, the *Annalen der Philosophie*, renamed it *Erkenntnis*, and converted it into the official organ of the movement. It is from this journal that most of the chapters in the present anthology are drawn.

The name 'positivism' was introduced nearly 150 years ago by Auguste Comte—a teacher of mathematics at the *École Polytechnique* in Paris—to designate a system of philosophy which was to be based on the 'positive method' (i.e., on the principle that human knowledge is restricted to observable phenomena). Its aim was to unite all phenomena under one vast empirical generalization. The several sciences were to be re-organized as branches of a single science, and arranged in a progressive 'hierarchy', with Mathematics and Physics at the base of the pyramid, and Sociology (a science founded and

named by Comte himself) and the 'Science of Morals' at the top. Metaphysics, like Theology, was to be thrown on the scrap-heap as a relic of a stage of thought which the human race had now outgrown; and Psychology, which Comte ridiculed as a 'metaphysical delusion', was to be replaced by Phrenology, 'a genuine physical, not a spurious mental science' (*Cours de philosophie positive*, 5 Vol. 1830-42).

To these empirical proposals, duly refurbished, the Viennese Circle added the golden rule that "all proof must proceed in logical order".¹ But the logic to be adopted was no longer the old scholastic system, which had lasted almost unchanged from Aristotle to the close of the nineteenth century. "The desire to substitute rigorous scientific methods for metaphysical poetry would remain a pious hope", says Carnap, "were no other instrument available: but fortunately a new weapon has been found", namely, what he calls *die Neue Logik*, developed by mathematicians like Russell and Whitehead on the foundations laid by the earlier work of Boole, Frege, and Peano. Besides the simpler 'one-place predicates' which formed the sole basis of the traditional subject-and-attribute deductions, the new logic recognizes 'two-or more-place predicates', i.e., relations; and this enables it to formalize the types of proof required alike in mathematics and in physics. Logical rigour is to be secured by the use of a symbolic notation. And according to Carnap all scientific concepts are to be reduced to terms of a 'physical thing-language', i.e., to the 'observable properties of things', where 'thing' means 'whatever occupies a region of the four dimensional space-time continuum', and the 'observable properties' are expressed by such predicates as 'red', 'large', 'heavy', 'hot'. Alternatively, says Carnap, we may speak of a reduction to terms of 'sense-data'.²

This twofold aim—the combination of logical rigour with a positive basis—was accordingly described as 'logical positivism', or, as some of its advocates preferred to say, 'methodological empiricism'. Their *credo*, as it developed, was characterized by three main articles of faith. (i) The 'positive method' was taken to require a wholly empirical criterion of meaning; and it was therefore held that whatever is incapable of verification is *ipso facto* meaningless. (ii) From this it was argued that, since the concepts of traditional metaphysics are all unverifiable, metaphysics can be wholly eliminated from philosophy: its problems have no solution because they are merely pseudo-problems. (iii) Philosophy is thus reduced to a single science; and, because the only verifiable concepts are those relating directly or indirectly to physical things, this 'unified' science is of necessity a physical science.

¹ On the other hand there is one important feature in the original 'positive method' on which Comte himself laid the greatest stress and which the Viennese Circle has unfortunately dropped—the implications of a *historical* view of the development of human thought and knowledge, interpreted in the light of the cultural background at each stage.

² See more particularly R. Carnap, *Der logische Aufbau der Welt* (1928) and his more recent *Introduction to Symbolic Logic and its Applications* (1957)—oddly enough, not included in the admirable bibliography, running to 65 pages, which is appended to the present volume.

The 'Circle' as such came to a tragic end shortly before World War II. In 1936 Schlick was murdered by a demented student on the steps of the University; Neurath had got involved in the Spartacist revolution in Munich; and the sceptical outlook of the other members aroused the hostility of Dolfuss, Schuschnigg, and the Austrian government. Finally, the advent of the Nazis and the outbreak of war forced nearly all the surviving members into exile.

Meanwhile, the movement had excited keen interest both in this country and in the United States. Karl Popper and later Ayer had attended many of the meetings. Ayer's brilliant book on *Language, Truth, and Logic* did much to popularize the main tenets; while Popper's *Logik der Forschung* and his later lectures did much to familiarize British students, often by way of criticism rather than support, with the new doctrines. In varying degrees Susan Stebbing, Ryle, Braithwaite, Wisdom, and Austin have all been influenced by the positivist standpoint. But in the main, I fancy, the method and doctrines of the British philosophers owe more to the contemporary 'linguistic analysis' which evolved out of the teaching of Wittgenstein, Russell, and G. E. Moore. And, although all of them have at one time or another been classed as 'logical positivists'—particularly by their critics—they themselves would prefer (as Professor Ayer protested) "not to be tarred with that particular brush".

Today those who adhere most closely to the positivist ideal would seem to be American logicians like Nagel, Morris, and Quine, whose views have brought them into close alliance with the American behaviourists, and a few of the younger American psychologists, like Skinner, Kessen, and Mandler, who have partly followed their lead. Much the same ideal commands the allegiance of several eminent philosophers in Poland, Holland and Belgium, in Norway and Sweden, and in Canada and Australia. Elsewhere—particularly in France, Italy, Switzerland, and present-day Germany—it finds hardly any support.

In compiling the present volume, Professor Ayer explains, an endeavour has been made to include papers illustrating the historical development of the movement, the range of its interests, and the main points of controversy. The chief contributors are the leaders of the original group; and some, notably Carnap and Hempel, have added brief appendices indicating how their views have since been modified. A few papers have also been added from writers who were not actual members—the famous article by Russell on 'Logical Atomism', one by Ayer on 'Verification and Experience', and another by Ryle on 'Philosophical Arguments'. To psychologists perhaps the most interesting part of the whole book will be Ayer's own editorial introduction. In what follows, however, I shall confine myself mainly to those contributions which bear on the problems of mind and consciousness; and I shall not hesitate to touch on the views of both earlier and later writers, particularly in this country, who have expounded somewhat similar doctrines.

Logical Positivism and Statistical Psychology. In Britain the philosophy of empiricism has had a far longer history than on the continent. It reaches back,

behind Hume and Hobbes, to the early Oxford nominalists, such as Occam and his followers. Among modern writers it might, I fancy, be plausibly claimed that the first 'logical positivist' was Karl Pearson, the 'father of British statistics'. Certainly Pearson anticipated many of the most distinctive tenets of the Circle. His famous *Grammar of Science* (1892), which had a profound influence on British psychologists at the beginning of the present century, opens with a reasoned appeal for a strictly 'positive' treatment of the whole problem of knowledge, and the ruthless expulsion of the 'metaphysical concepts' which still encumbered the science of his day. "The metaphysician is a poet," he says, "and often a great one; but he is not recognized as a poet, and therefore becomes a danger to the community." Special emphasis is placed on the need for factual verification at every stage of the analysis. "To Comte himself", so Pearson declares, "science owes a debt of gratitude: for Comte repudiated all ontological hypotheses, and taught that the sole basis of knowledge is experience." But it is from Mach rather than from Comte that he derives the most characteristic of his 'anti-metaphysical' doctrines, particularly his principle of 'economy of hypotheses'. The later portions of his work consist of a plea for a more rigorous methodology: in his view it is not so much the *results* of physical science that call for reconsideration as the *language* in which they are expressed. The 'grammar' he proposes is partly qualitative. This consists of a modernized logic which is to be hypothetico-deductive rather than 'inductive' in the sense of Bacon or Mill. But wherever possible quantitative methods are to be applied. These are to be based on the mathematical theory of probability, which he himself derived mainly from the work of Gauss and Laplace. In the formulation of scientific laws the notion of 'correlation' is to be substituted for the traditional concept of 'cause'. Comte, it may be remembered, had dismissed the idea of applying the calculus of probability to the solution of scientific problems as "purement chimérique, et, par conséquent, tout à fait vicieuse"; and Mill's chapters on the subject were far behind the mathematical teaching of his day. On the other hand, Pearson believed that, supplemented by the new statistical techniques he himself had devised, a logic based on probability would give the biological or 'organic' sciences (including anthropology, psychology, and sociology) the same exactitude as was possessed by the physical or 'inorganic' sciences; and so in the end bring all branches of knowledge under the aegis of one all embracing 'synoptic science'.

The Viennese Circle appears to have included no biologists; but it numbered among its members and earliest associates mathematicians, as well as logicians, of considerable originality. The views on probability which they elaborated differed widely among themselves; and it was to a large extent the ensuing controversies that led to the development of a greatly improved basis for statistical theory. (1) According to Reichenbach no proposition is or can be certain. Hence truth and falsehood must be abandoned as absolute concepts; they merely designate ideal and limiting cases of a continuous series of

probability-values. Logic therefore becomes essentially a calculus of probabilities ; and on this foundation a general theory of induction is built up, with the aid of certain extra-logical ' posits '. (2) The theory of von Mises comes nearer to that with which British psychologists are familiar from the writings of Fisher. It is essentially a frequency theory, and limits probability to those inquiries which are concerned with a *Kollektiv* (" a series of observations, arising either from mass-phenomena or from some kind of repetitive process, producing a series which can be conceived as being extended without limit "). This permits us to measure probability by the ratio of the number of two classes ; and together with the wholesale rejection of Laplace's doctrine of inverse probability yields a statistical theory which, as its advocates claim, is thoroughly empirical and positivistic. (3) The official view of the orthodox positivists is based on that worked out by Wittgenstein and Waismann. " Logic ", says Wittgenstein, " is the investigation of regularity ; outside logic all is accident " : the two spheres are wholly distinct. There is no such thing as a logic of induction. Generalization proceeds by the construction of truth-functions of elementary propositions with the aid of a symbolic matrix called a ' truth-table '. Probability is neither objective, in the sense that it can be predicated of events (for an event will either occur or not occur : *tertium non datur*) ; nor is it purely subjective, in the sense that it merely affords a measure of our own belief as limited by the limits of imperfect knowledge. A third interpretation is accordingly proposed. Probability is defined as " the logical relation between two propositions " : its amount expresses the degree of overlapping between the ' fields ' (*Spielräume*) of the two propositions as determined by the number of its ' truth-grounds '. This conception and much of the subsequent development closely resemble the treatment adopted by Keynes. On the other hand, unlike Keynes' theory, it is applicable only where a metric is possible.

One thing these three-cornered discussions have made clear : measurement and statistical methods generally are valid only where certain specific conditions can be postulated—a requirement that the statistical psychologist too often omits to check. But in many ways the most suggestive outcome of the prolonged and somewhat inconclusive controversies has been the slow and unexpected discovery that " empiricism as a theory of knowledge has proved inadequate, though less inadequate perhaps than any previous theory." It appears that the methodology of the positivists can be justified only if the universe possesses certain definite characteristics which are neither logically necessary nor themselves capable of justification on empirical grounds. In short, as Russell puts it, " it is an inescapable conclusion from the logic of probability that scientific inference requires for its validity *principles which experience cannot even render probable* " ¹.

To the general psychologist the most interesting feature in the whole doctrine of logical positivism is the strong support it seems to lend to the theories

¹ *Human Knowledge : its Scope and Limits*, 1948, p. 13, cf. p. 527.

of the behaviourists. Otto Neurath has argued at some length that, since 'psyche' and 'psychical' are, by nature and by history, just 'metaphysical' concepts, what is commonly called 'psychology' should be replaced by 'behaviouristics'. The traditional terms of the psychologist's vocabulary—mind, will, emotion, intelligence, etc.—are all just fancy-names for what Hahn has dubbed 'überflüssige Wesenheiten'. In this country Gilbert Ryle in his well-known book on *The Concept of Mind* has taken up a similar position. "Consciousness", he tells us, "as ordinarily understood, is a myth." The bifurcation of behaviour into overt physical processes and occult mental processes—as though the cerebral machine held ensconced within itself a lurking ghost—is the consequence of a 'category-mistake'. Self-observation is impossible, because there is no self to observe. And psychologists are urged to abandon once for all "the notion that 'psychology' is the name of a unitary inquiry". Thus, as one sympathizer has already remarked, "the old positivism dealt psychology a mortal blow, and the new positivism has buried it".

It is unfortunately the statistical psychologists who seem to have taken these pronouncements most seriously, as may be readily seen from the contributions and quotations in the pages of this journal. In the early days of statistical psychology no investigator would have carried out an experiment, or have attempted to apply a newly constructed test, without securing introspections from his subjects in order to ascertain the actual nature of the mental processes which his experiments or his tests were eliciting. Now he can save himself the trouble, for both positivists and behaviourists have conclusively shown (at least so he believes) that introspection is either self-contradictory or self-defeating. This therefore is a point we must take up in a moment.

On the other hand, the insistence of the logical positivists on the need for a more rigorous methodology has proved both wholesome and well timed, for of all the natural sciences modern psychology was until recently the least methodological. Watson's earlier behaviourism was based on a series of dogmatic and provocative assertions, which his own experimental work in no way established. Hull's new behaviourism rests squarely on the axiomatic method, as developed by the later logical positivists¹: "eighty-six definitions and eighteen postulates", he announces, "have been formulated in the notation of symbolic logic"; and from these are to be deduced a systematic set of theorems to be verified empirically by suitably planned experiments: that at any rate, so he says was the goal at which he aimed. Similarly Thurstone's theory of multiple factor analysis, Bush's and Mosteller's 'stochastic models' for the learning-process (to cite but two of the many later examples) both strive to set out their arguments in a lucid and logical fashion which was almost unknown

¹ Actually the use of symbolic logic was (as Hull relates in his autobiography) indirectly derived during a spell of joint work with Woodger just after Woodger himself had published his *Axiomatic Method in Biology* (1937). Woodger's symbolism was derived mainly from the *Principia Mathematica*; but his ultimate aim was "to eliminate metaphysical elements from biological topics", and he acknowledges special help from Carnap and Tarski.

in the days of Spearman and Freud. Piaget's studies of *l'épistémologie génétique*, as he himself explains, were largely inspired by the 'logistics' of the positivist school. And finally Mandler and Kessen in their recent volume have made a systematic effort to apply the principles of Carnap and his American followers to the construction of a 'psychological language' that shall be 'genuinely scientific'¹.

Among British psychologists there has never been quite the same interest in formal logic and the use of symbolic techniques that has from time to time been displayed by certain American theorists². Those who have advocated an improved methodology have more commonly fallen back on the old notion of a 'hypothetico-deductive procedure', supplemented by occasional catch-words from operationism, information-theory, communication-theory, and the like, without any very noticeable advances in their own logical techniques. And where British textbooks have favoured a predominantly physicalistic approach to psychological problems, that is due rather to the American behaviourists than to the positivists of this country or the continent. Nevertheless, this seemingly lax eclecticism has not been without certain compensatory merits. It has fostered a broader outlook; it has kept theorists in touch with everyday phenomena as well as with laboratory results; and it has allowed both methods and conclusions to be judged, not by some doctrinaire criterion, but at the bar of commonsense.

Verifiability as the Criterion of Meaning. In their attacks upon psychology the argument adopted by the modern positivists is not so much that there is no such entity as a 'mind', a 'soul', or an individual 'consciousness', but rather that these very words are meaningless. But what do they mean by meaning? In order to determine whether or not a concept is meaningful they have brought forward a new criterion which is applied with devastating effect in almost every chapter of this book. As originally formulated, the principle ran as follows: "A sentence has a meaning only if it is either (a) analytic or (b) capable of conclusive verification at least in principle by observational evidence." How far can we, as psychologists, subscribe to such a criterion?

To begin with, it knocks the meaning out of every universal statement ("All men are mortal", "All metals are good conductors", and the like); for every empirical (i.e., non-analytic) proposition which is cast in this general form is essentially a hypothesis, and as such continually subject to revision by later experiences; consequently no such statement can ever be conclusively verified. "To Generalize", said William Blake, "is to be an Idiot." And

¹ G. Mandler and W. Kessen, *The Language of Psychology* (1959).

² The notation of symbolic logic has at times proved extremely useful, not only in constructing tests of intelligence and other mental abilities, but also in analyzing the development of reasoning in young children and the typical thought processes of educated and uneducated adults: (see, for instance, D. Wheeler, 'Studies in the Development of Reasoning in School Children' this *Journal*, XI, pp. 137-160 and refs.).

Schlick does not hesitate to declare that all generalizations are pieces of nonsense, albeit at times quite important pieces of nonsense. To avoid such paradoxes Ayer has proposed a 'weaker' sense for the term 'verifiable': "in the weak sense a proposition is verifiable if it is possible for experience to render it *probable*". Even this version he now would prefer to amend. "A statement", he suggests, "is directly verifiable if it is either itself an observation-statement or is such that, in conjunction with one or more observation-statements, it entails at least one observation-statement which would not have been deducible from these other premises alone." An analogous and slightly more elaborate definition is appended to cover cases of 'indirect verifiability'.

At this point the reader will be tempted to inquire: what is the meaning of Ayer's own formula? Is the formula itself capable of verification? If not, are we not obliged, in sheer consistency, to condemn it as meaningless? Plainly it is not a mere analytic proposition. Are we then to regard it as describing the way philosophical disputants actually frame their concepts and their statements? And if so, may we not try to verify it by observing how they do reason? The answer is tacitly given by the logical positivists themselves: for it is their constant complaint that philosophers have hitherto failed to apply the correct criterion.

It seems clear, however, that their principle is intended to be, not descriptive, but prescriptive. And Professor Ayer has recently maintained that we should regard it as a definition. Nevertheless, as a definition it would seem to be decidedly arbitrary. One might almost imagine it had been ingeniously framed so that those who dislike certain metaphysical or psychological doctrines should have a reasonable ground for rejecting them. At all events it certainly does not define the meaning of 'meaning', as that term is ordinarily employed. Suppose someone says: "The Creator of the Universe is good." The logical positivist will wave it aside at once, with the retort: "There is no need to examine the evidence for a creator, or to adduce the presence of evil to show that, if he existed, he could not possibly be good, for the assertion is quite unverifiable, and is therefore devoid of meaning." The less sophisticated listener, however, will surely retort: "I know the meaning of the words 'Good', 'Universe' 'Creator': and consequently I understand the meaning of the statement as a whole, although I readily admit that, even if true, it would be quite impossible to verify it."

That being so, it is tempting to argue (as Carl Hempel does in the present volume) that, for a statement to be meaningful, it is both necessary and sufficient if its non-logical *constituents* refer, directly or indirectly, to observables, whether or not the statement as a whole is capable of verification. This alternative criterion would still rule out such psychological concepts as the 'ego', i.e., the 'subject' implied in the famous "*cogito, ergo sum*": for the 'subject' is neither observable nor reducible to terms of observables. As Hume long ago remarked, "When I enter most intimately into what I call *myself*, I always stumble on some particular perception . . . I never catch *myself*, and never can

observe anything but a perception." Plausible as it is, however, a little reflection will show that the definition would be equally fatal to many of the more comprehensive theories of modern physics. No one has ever observed the lines of force in an electric field. No one has observed, or could observe, a proton or an electron. It used to be said that "the ether is just the nominative of the verb to undulate"; and it might equally well be said that "the electron is just the nominative to the verb to jump-from-orbit-to-orbit-and-elsewhere". But even the jumps and the orbits are unobservable—unobservable in principle as well as in fact. Electrons can be treated either as particles or as waves. As Eddington once put it, we might just as well say: "eight slithy toves do gyre and gimble in the oxygen wabe; and only seven in the the nitrogen". The name itself is no more than a symbol. In Heisenberg's formulation, for example, it simply stands for an array of numbers manipulated according to the rules of matrix algebra. In short, what the physicist affirms are merely certain relations or a certain structure consisting of relations: the *relata* themselves are wholly unknown. The theories of modern science describe conceptual models, not portions of observable reality. They start with primitive and undefined concepts and primitive and unproved propositions: and it is rather the corollaries of the system that we seek to verify than the concepts or postulates themselves¹.

Physicalism and Psychology. Let us now examine how precisely the logical positivist applies his new criterion when he turns to discuss the problems of the psychologist. The clearest pronouncements are to be found in Carnap's paper on 'Psychology in Physical Language'. "All sentences of psychology", he maintains, "describe physical occurrences, viz., the physical behaviour of humans and other animals": only in that way can they acquire meaning. Here the ghost of Comte is clearly hovering in the background. Comte, it will be remembered, sought to reduce psychology to a branch of physiology and so in the end to a section of "the one unified science which is based on physics". Nevertheless, there are marked differences between Comte and Carnap. Comte, like the modern behaviourist, held that introspection was not a method that the scientist adopts, "since it is as impossible for a self to observe itself as it is for a cannonball to hit itself". Carnap, on the other hand, denies that "the psychologist who adopts the method of introspection thereby exposes himself to criticism . . . We may apply any *method* we choose: we cannot, however, *interpret* the sentences so obtained just as we choose".

How then does Carnap prove his fundamental thesis? In dealing with 'Sentences about One's Own Mind', he takes as an illustrative instance the

¹ Cf. C. Burt, 'Definition and Scientific Method in Psychology', this *Journal*, XI, 1958, pp. 37 f. Hempel too, in the article reprinted in the present volume, makes similar reservations: "what is sweepingly referred to as 'the meaning' of a given scientific hypothesis cannot" (he says) "be adequately characterized in terms of potential observational evidence alone; the statements of empirical science have a surplus meaning over and above what is expressed in terms of observation sentences."

statement "I am now excited" (P_1). The 'protocol sentences', p_1 , which rationally support this statement (he says) have some such form as "I feel my hands trembling", "I hear my voice quavering", etc. ; and P_1 "subsumes all the possible sentences of this sort". It follows that P_1 has the same content as the sentence P_2 : "My body is now in that condition which, both under my own observation and that of others, exhibits such and such characteristics of excitement" ; and this is a purely 'physical sentence'.

Now much of the plausibility of this argument arises from the kind of example which Carnap has selected. Being excited is a condition which is as much physical as mental. Still it seems a little quaint that I myself should have to 'support' my impression that I am excited by observing that I feel my hands trembling, etc. But what the critic naturally asks is—how can *Carnap* consistently make statements about this latter feeling which after all is mine and not his ; and if he can accept my statement about feeling my hands trembling as meaningful, why can he not (as the unsophisticated person would surely do) accept my unsupported statement that I feel excited ? His point apparently is that the trembling can be observed by others, and is therefore a 'physical' phenomenon. But most adults would have sufficient self-control to check any visible movement ; and, if it occurred, would usually fail to notice it. Hence in practice a psychologist is almost bound to accept his own and his subjects' introspective statements without further support, or else reject them on the ground of their intrinsic inaccuracy.

But the issue involved would become much plainer if we chose a more representative instance of conscious experience—something that has no bodily expression. Looking up at the white ceiling, I say : "I see a dark round blob". My companion does not retort : "Since I cannot verify your alleged sensations, I don't understand your meaning." He may possibly remark that he himself sees no such blob. But even so he would not dismiss my 'sentence' as meaningless. Rather he might reply : "Yes, but isn't that a hallucinatory sensation ? And shouldn't you ask your oculist about it ?" But suppose he happens to look out of the window at the sun as I had done : then afterwards he too will observe a similar blob. To that extent he *can* verify my statement. But he verifies it, not because the thing is physical, nor yet because he perceives a sensation which is numerically the same as mine. He verifies it by performing an act of introspection for himself. The logical positivist, however, protests that "since on your own admission he does not see *your* sensation, he does not really verify *your* assertion". Apparently the only kind of verification that he will accept must be based on perfect identity. My companion must actually be able to gaze inside my brain—science or telepathic skill having first devised the means of obviating the thickness of my skull and the turbidity of its contents—and there he must have a direct perception of *my* sensation : only then will our positivist critic be entirely satisfied. But in no other branch of science is so strict a requirement imposed on the observer. If I declare that "This piece of ice melts at 5° C, not 0° C", my fellow-scientist does not reply that, because

he cannot also melt exactly the same piece, he cannot verify my assertion, and therefore rejects it, and all others like it.

But there is a still more potent objection. In cases like those I have envisaged—the observations which led to the discovery of after-sensations, of colour-mixture, of subjective visual experiences like the ‘fortification-patterns’ which often herald an attack of migraine, and the wide individual differences in mental imagery—it would be quite impossible to translate the observer’s descriptions into ‘physical sentences’ of the type which Carnap demands¹. And yet, had it not been for introspective observations of this kind, psychologists would never have recorded these and many other highly instructive phenomena. If for the future psychologists decline to carry out introspective inquiries or to accept introspective statements at their face value, many new contributions to knowledge will almost certainly be lost.

Carnap’s further discussion, however, might seem at first sight to provide him with a possible loophole. When the sentence P_2 refers to the ‘characteristics of excitement’, he explains that these include, not only those characteristics referred to in p_1 , but also those mentioned in protocol sentences of another type, p_2 , which actually have the same verbal form as P_1 itself—e.g., the protocol sentence “I now am excited”—or, as he would prefer to phrase it, “Excitement now”. But he at once goes on to insist that this must also be “physically interpreted”: otherwise “A’s protocol sentence p_2 could not be tested by B, and would thus be meaningless to B”. However, as we have just seen, this interpretation will in many cases be quite impossible. Carnap’s solution, it would seem, is to require us to undertake a sort of book-keeping by double-entry: but the difficulty then is to make the two accounts balance.

At this point we encounter yet another problem which we have still to consider, namely, what, if anything, B can legitimately assert about the consciousness of A? And this brings us to the second part of Carnap’s main argument—that which relates to “Sentences about Other Minds”. Much the same treatment is adopted as before. Carnap takes the sentence “Mr. A is now excited” (P_1), and argues that a sentence like P_1 “has the same content as a sentence P_2 which asserts the existence of a physical structure characterized by the disposition to react in a specific manner to specific physical stimuli” (i.e., in the present case, Mr. A’s physical organism, which is now liable to react to questions, etc., with accelerated breathing, confused replies, etc.). As for the excitement itself, or the consciousness of which it is an attribute, that we are told is “an occult property which stands behind physical structure, but in itself is unknowable: if the content of P_1 went beyond that of P_2 (by asserting

¹ Someone might contend that what the observers are really reporting are minute changes in the physical structure of the retina or brain. But Carnap expressly says that his equivalent ‘physical sentences’ are not intended to be ‘statements in terms of micro-physiology, since these, as a rule, are not known at the present stage of scientific development’. To treat the statements themselves as ‘verbal behaviour’ and therefore ‘physical’ (regardless of their content or significance), as some writers have suggested, seems a rather desperate way of ‘saving appearances’.

consciousness), then the additional component would not be testable, and would therefore be meaningless".

Carnap concludes by pointing out that "in its broad outlines the position here advocated coincides with the psychological position known as 'behaviourism'" —at least in respect of its epistemological principles, if not of its special methods. And he then goes on to cite and rebut the commonest types of objection. "If", asks the critic, "A, whom I know to be a truthful and accurate observer, assures me that he has a feeling of excitement, should I not accept his statement as at least probably true regardless of any physical reinterpretation?" Carnap's answer is simply to insist once again on his criterion of meaning. Before B can accept A's statement, he must first be able to understand it: but if it is not to be interpreted physically, he cannot test it; and it must be meaningless for him. However, for a critic who declines to accept this new Pickwickian definition of meaning, Carnap's rebuttal can carry no conviction. As everyday experience shows, except for behaviourists and logical positivists in their official moments, all of us regularly accept other people's testimony about their conscious states as perfectly intelligible and perfectly credible. No modern psychiatrist says to his patient, "You assure me that you have a pain in your arm; yet you display no physical symptoms of that pain, and I can find nothing physical to account for it. Therefore I must dismiss your statement as meaningless."

The second objection is based on the well-known appeal to analogy. "When I myself am angry" (so Carnap represents his opponent as saying) "I not only act the behaviour-pattern of an angry man; I experience a *feeling* of anger. Accordingly, when I observe someone else behaving in the same fashion, I conclude that he too has a feeling of anger, which is something more than a merely physical state." Carnap's answer is to deny that we have here the conditions necessary for a valid analogy. By way of comparison he takes first a perfectly legitimate example. Yesterday B found an oblong box with a yellow label, and discovered that it was filled with matches. Today A shows him a box of the same size, shape, and colour; and B consequently infers that A's box in all probability also contains matches. Here A's box is verifiably similar, in all relevant respects, to B's box. The inference is therefore valid because in both premiss and conclusion we can simply substitute A for B. But in the case of conscious experiences the requisite conditions no longer obtain. To begin with, B could never directly compare A's feeling of anger with his own, as he could, by way of a check, compare A's matches with his own; and secondly, when B describes his own feelings, he is not really asserting a specific property of a specific entity, as he does when he says: "My box contains matches." All he can legitimately say, according to Carnap, is: "Now anger", or "Anger occurs." Here therefore it is no longer possible to substitute A for B.

I find neither of these arguments convincing. Let me take the second point first. "Anger occurs" is a very inadequate protocol sentence. When

and where does the event occur? B should at least report : " 2 a.m., April 1, 1960, in B's bedroom at such an address, *Anger occurs*." Even if we hold that such a sentence as " I, B, feel angry " implies an unverifiable ' subject ', B's name must be mentioned somewhere. And the same should hold of A. Moreover, anger is not an event that occurs in complete isolation, as though a feeling suddenly exploded in a vacuum. It emerges as partial change in a vast continuum, the specific continuum we call so and so's consciousness. Consequently, in my view, the sentence " I am angry " is still the better formula ; and we can surely interpret it, not as ascribing an introspective attribute to an unknown entity, but as stating that anger occurs ' now ', at a specified date, as an incident in what the speaker would refer to as ' *my* life-history '. And for this purpose I should be content to construe ' my ' as a word expressing the relation of that life-history to a particular physical organism, ticketed for purposes of identification with the monomark B. If so, we can quite validly substitute, in both premiss and conclusion of the analogical inference, the name A for the name B.

As for Carnap's other point, I should reply that, when B remarks " A is angry " (i.e., as I should interpret it, " Anger is occurring as part of A's life-history ") what he is really asserting is, not a directly observable fact (like A's overt behaviour), but a hypothesis. Now in science we proceed for more often by disproving hypotheses than by verifying them. If a hypothesis, unverifiable in itself, nevertheless affords the simplest and the most satisfactory explanation of a given type of situation, then, provided there is no factual evidence to refute it, we are, as a rule, willing to adopt it. Accordingly, if A says " I see a dark round blob ", I do not scruple to accept his statement at its face value. When he says " I see a green field ", I tacitly assume that the colour he sees is virtually the same as I see, unless I already know he is colour-blind. However, in most cases what we are really concerned with are not the colours and the qualitative contents of other people's consciousnesses, but the relational structure of their experienced life-history¹ ; and here, from their behaviour and conversation we have every reason to infer that, for persons of similar sex, age, upbringing, and the like, the relational structure is basically the same ; when it differs, it is frequently possible for the psychologist to say where and how it differs.

After all this is the attitude every one of us adopts in our day-to-day lives. In his official moments the logical positivist may perhaps persuade himself that for him only a man's outward behaviour, and not his inner consciousness, can possess any assignable meaning. For the novelist, the biographer, the teacher, the clinician, and above all the student of individual psychology, a

¹ Schlick maintains that this is the only type of case in which the argument from analogy is justified. I can, he maintains, legitimately compare two spots of colour in my *own*-consciousness as similar, but not two spots in two *different* consciousnesses : hence, when we are concerned with the experiences of two persons, the term similarity must be redefined so as to refer only to similarity of order or structure, not to similarity in quality. Ayer, however, would reject this limitation : " there is," he declares, " no warrant here for separating structure from content and for arguing that structure can be communicated, whereas content cannot ". I should certainly agree. (See A. J. Ayer, *The Problem of Knowledge*, 1950, p. 208, and his arguments in support of his view.)

sympathetic understanding of the mental life of others is not only possible, but indispensable. I have no wish to extol the intuitive and impressionistic methods of a *Verstehende Psychologie* in sharp opposition to the more critical and analytic methods of the physical sciences. Nevertheless, both from my personal experience and from my work in training others, I am convinced that one of the most valuable assets a child psychologist can possess is the ability *sich hineinzusetzen und hineinzu fühlen*—to place himself within the mind of the child he is studying, feeling his emotions, thinking his thoughts, and temporarily stocking his own mind with the kind of percepts, memories, and mental images which the child possesses. This means not only that he must familiarize himself with the child's environment and watch with an eagle eye the child's outward behaviour : it also means that he must have a lively interest in other persons, and above all a keen, imaginative insight—a sort of *Empathy*, which may in part be acquired by practice, but is to a large extent the gift of a particular type of temperament. Certainly there are some students of psychology who appear quite powerless to achieve it.

Logical Difficulties. In addition to these psychological objections there seem also to be certain logical inconsistencies inherent in the arguments which Carnap has brought forward.

(1) On his own assumptions the distinction between conscious experience and overt behaviour can be drawn by Carnap only in reference to himself. If Carnap were just soliloquizing—telling himself that, “The idea that I have thoughts and feelings would be meaningless to other people, so I must translate them into words or gestures”, then his arguments would be perfectly consistent. But he wishes to extend his advice to us. Yet, since the idea that we may have thoughts and feelings must on his own showing be meaningless to *him*, how can he logically do so? His attitude seems to be that of the child in the story : “Say ‘Thank you’.” “I can’t.” “What can’t you do?” “I can’t say ‘Thank you’.”

(2) Even if we waive this objection, any attempt to communicate our own experiences to one another must still, with Carnap's interpretation, always leave us at cross purposes. My dentist asks whether I still suffer toothache, since my teeth seem perfectly sound. According to Carnap, his question must relate to my physical condition, for any reference on my part to my own sensations would be meaningless to him. How then am I to reply? If I say I sometimes feel an unpleasant twinge, I am not answering the question Carnap says he has put ; if I do not mention these feelings, I am in effect telling a lie. Worse, still, on Carnap's assumptions, I cannot think about *my own* past feelings. Being past, they must now be completely unverifiable, and hence for me completely meaningless. Carnap, of course, is not unaware of these logical difficulties, and has made heroic efforts to meet them. But even his most subtle and sophisticated restatements have not succeeded in providing an escape from the linguistic solipsism to which his initial assumptions commit him.

(3) Finally, the line of demarcation which physicalists and behaviourists attempt to lay down between what is public and what is private is bound to be both arbitrary and inconstant. On their own showing B's perceptions of his own body are peculiar to him ; and A's perceptions of B's body must be peculiar to A. B's perceptions of it are mainly organic ; A's mainly visual ; and there are no elements common to both. The observations of B's body must therefore be private to each one. Nevertheless, according to the behaviouristic view, B's body ranks as a public object, and can therefore be referred to in public language. Plainly these two contentions are at variance with each other.

When for the sake of argument it becomes convenient to contrast what is public with what is private we tend to draw the line differently at different times and for different purposes. For example, before it became customary to distinguish secondary qualities from primary, such things as sound, colour, hardness, and the like would have been unhesitatingly classed as physical properties, and in that sense 'public'. However, as a consequence of certain intersensorial correspondences, it appeared that most of these secondary qualities could be explained in terms of the primary, while the primary in their turn could be broadly identified with what we ordinarily think of as material characteristics—characteristics which we tend to visualize as spatial structures. On these grounds all secondary qualities are now usually held to be private, whereas structures residing in our common three-dimensional space are regarded as public. But from the standpoint of sub-atomic physics even the material properties of classical physics are now perceived to be highly subjective interpretations of mass-phenomena, interpreted at a macroscopic, i.e., a human, level. At the same time, both quantum theory and the theory of relativity have now deprived the present-day physicist of his cherished illusion that the universe, or parts of it, can be adequately described with no reference whatever to an observer.

Moreover, just as most of the things that the behaviourist alleged to be public now turn out to be (in one sense, at any rate) private, so what he commonly regards as private may on occasion be fairly included in what is public. Ordinarily we should agree that our feelings are private. Yet, at times, when, as we say, 'a wave of feeling' sweeps over a whole nation, we are quite ready to talk of 'public feeling'. Indeed, there seems no cogent reason why the vocabulary of the public language should be restricted solely to so-called public objects : nor why the logical positivist should insist on identifying what he calls the 'universal intersubjective language' with a language that is exclusively 'physical'.

In nearly all these arguments, it will be seen, everything in the end hinges on the 'verifiability criterion'. It is applied at every turn. It is used to sweep away many time-honoured notions common to the plain man and the philosopher. It provides a ready means for removing at the very outset all those troublesome concepts of introspective psychology which have hitherto defied every attempt at translation into purely physical terms and have thus continually entangled

the forward march of the out-and-out materialist. In short, it would seem, the standard method of the logical positivist is just that of the traditional Greek hero in a hurry to conquer the world—cutting through every knot that his predecessors failed to untie.

Moderate Positivism. The force of the foregoing objections has been, in varying degrees, acknowledged by many of the positivists themselves. Schlick's views, which are represented by four papers included in the book which Ayer has edited, are far from maintaining the same uncompromising attitude as that of Carnap and Neurath; and several of the other contributors—notably Hempel, Stevenson, and Waismann—now evince a disposition to retreat from some of the more paradoxical formulations put forward in the earlier stages of the movement. With nearly all of them the stringency of the 'verifiability criterion' has been greatly relaxed; by some it is replaced by a mere 'confirmability' criterion; 'emotive' meaning is recognized as well as 'cognitive' meaning; and semantic theories are introduced to supplement a purely syntactical analysis. Several of the more recent representatives of the school are even willing to allow that the traditional problems of philosophy are not in every case to be eliminated by a more suitable choice of linguistic forms; and ontological issues are now often discussed in and for themselves. Thus, as Ayer puts it, the metaphysician is no longer treated as a criminal to be summarily despatched but rather as a patient to be reasoned with, or at any rate humoured.

Ayer's own views on the psychologist's problems are most clearly set forth in his book on *Language, Truth and Logic*. Here it is somewhat surprising to find him declaring that several of the questions we have examined above are to be solved, not by empirical evidence, but by *a priori* argument. "The answer to the question whether sense-contents are mental or physical is", we are told, "that they are neither . . . The distinction between what is mental and what is physical does not apply to sense-contents: it applies only to objects which are logical constructions out of them." The self is not a 'metaphysical entity', but a logical construction out of sense experiences; and, once we are "freed from metaphysics, we see that there can be no *a priori* objections to the existence either of causal or of epistemological connections between minds and material things".

An important corollary is drawn. In order that two or more sense-experiences can belong to the sense-history of the same self "it is necessary and sufficient that they should contain organic sense-contents which are elements of the same body"¹. Now it is "logically impossible for any organic sense-

¹ It is a little surprising to find sense-contents described as 'elements' of a physical body. This apparently means that the relation between my sensations and my body is a relation of part and whole. If we accept this contention, we should naturally be led to conclude straight away that a man's consciousness cannot survive the dissolution of his body any more than the legs and head of a snowman can survive when all the snow has melted. Elsewhere, however, Professor Ayer remarks that "we do not hold (as Hume did) that the self is an aggregate of sense-experience, or that the sense-experiences which constitute a particular self are in any sense parts of it"; so I may have misinterpreted him.

content to be an element of more than one *body*". Hence "belonging to the sense-history of the same self" (so it is argued) must denote a relation which is both symmetrical and transitive; and from this it follows that a sense-experience cannot belong to the sense-history of more than one *self*. If then personal identity can only be defined in terms of bodily identity, it becomes "self-contradictory to speak of a man as surviving the annihilation of his body".

On much the same grounds Ayer is here led to reject the argument from analogy¹. Since A's experiences cannot possibly form part of B's sense-history B cannot observe them. Hence the hypothesis which the argument from analogy sets out to justify is completely unverifiable; and "no hypothesis can render probable a completely unverifiable hypothesis". Here, however, the psychological reader should perhaps be cautioned against basing any hasty inferences on the result thus reached. As Ayer observes later on, "the conclusion to be drawn is, not that the existence of other people is for me a metaphysical, and so a fictitious, hypothesis, but that the assumption that other people's experiences are completely inaccessible to my observation is false". The introspectionist, therefore, even if he is unable to endorse every step in Ayer's reasoning, seems in the end to get nearly all the freedom he could wish.

Nevertheless, as a psychologist I cannot help feeling that many of the questions that Ayer here attempts to solve by *a priori* reasoning are of the type that can only be solved (if at all) by empirical inquiry. Many psychologists and philosophers (James, McDougall, McTaggart, Broad) have held that it is

¹ The argument from analogy has often found a place in psychological textbooks. Stout, for example, writes: "No one can directly observe what is passing in the mind of another: he can only interpret external signs *on the analogy* of his own experience" (*Manual of Psychology*, by G. F. Stout, revised by C. A. Mace, 1938, p. 45, cf. also pp. 583 f.). It is not, I think, always realized that the argument is meant to explain, not how, as a matter of psychology, we come to interpret the conscious experiences of others, but only how, as a matter of logic, we may justify our interpretations. Even so, it is by no means a convincing method of justification, particularly in the case of emotional states. As Lipps long ago pointed out, what I observe of my own bodily behaviour is quite different from what I observe in the case of others ('Das Wissen von fremden Ichen', *Psychol. Untersuch.*, I, pp. 4 f.). Carnap's example of anger is thus particularly unfortunate: when a man loses his temper, he is often the last person to notice it.

Our own recognition of other persons' emotional experiences starts in the earliest years of childhood, and seems (as McDougall suggests) to be the result of a kind of 'primitive sympathy' which is instinctive: the expression of pleasure or of fear by its mother directly excites a similar feeling in the child, and it is in this way that he comes to understand what she is feeling (*Social Psychology*, pp. 79 f. and 158). In general I should say that it is in the course of conative rather than of cognitive experiences—during a conflict of wills or through the emotion of love—that we come to realize that other persons are conscious beings just like ourselves. When we do fall back on analogies in seeking to fathom another person's mind, it is more often on some analogy derived from observation of others we previously have known rather than from observation of ourselves. So far as their cognitive states are concerned, our reasoning seems rather to be tacitly based on what older writers would have called 'induction'. Whenever I have looked at a bright light, such as the sun, I notice that I subsequently experience 'after-images' of the same shape. I inquire of others; and they report a similar effect. I therefore infer a kind of 'causal law'; and on the strength of that 'law' I predict to my students that, if they look at a bright light, they too will in all probability see an after-image. This kind of reasoning would at once be accepted as valid in other sciences; I cannot see why the behaviourists should reject it in the case of psychology.

at least conceivable that a person might survive the dissolution of his body ; and, unless we deliberately re-define the terms so as to exclude that possibility, it is difficult to accept the view that such a hypothesis is self-contradictory. Take, for instance, a concrete case. Messages purporting to come from someone named Mary were received over a period of many years by the wife of Professor Verrall, by Mrs. Fleming (the sister of Rudyard Kipling), by Mrs. Coombe-Tennant, and by several others ; and these, when pieced together, seemed to recall many long-forgotten events and circumstances, some known apparently only to Miss Lyttelton (daughter of Lord Cobham) who had died on Palm Sunday 1875 and to Lord Balfour, who had intended to propose to her. Lady Balfour (wife of Arthur Balfour's nephew, the present Earl) has recently published the relevant data, and discusses among other hypotheses (without explicitly endorsing it) the view that the ' messages ' emanate from a conscious and purposeful being who is to be identified with the person once known as Mary Lyttelton¹. The sceptical reader may quite legitimately question the accuracy of the records, the symbolic interpretations, the validity of singling out a series of suggestive coincidences from a large mass of material, and the claim that unconscious telepathy or unconscious memories on the part of the recipients were excluded by the circumstances of the case ; but he surely would not be justified in dismissing the hypothesis without further hearing on the ground that it involves a logical self-contradiction, any more than the adherents of classical physics in the 1920's were justified in dismissing the quantum-hypothesis because on the surface it seemed full of logical self-contradictions. Ayer's version of the survival-hypothesis only issues in self-contradiction because of the definitions he adopts for the crucial terms².

Nor should I agree that for any two or more sense-experiences to belong to the same self it is necessary and sufficient for them to contain " organic sense-contents which are elements of the same body ". There are moments when I seem wholly unaware of my body ; certainly, so far as introspection goes, my experiences at those moments contain no ' organic sense-contents '. Moreover, whether or not we believe in the occasional occurrence of multiple personalities (as in the Beauchamp case, for instance, where Sally, and the various Miss Beauchamps all mention organic sensations derived from the same body), it seems evident that such cases *might* occur. The real distinction between A's personality and B's surely rests on the peculiar unity and continuity exhibited

¹ The Countess of Balfour, ' The Palm Sunday Case : New Light on an Old Love Story ', *Proc. Soc. Psychic. Res.*, LII, 1960, pp. 79-267 : in point of fact, the alleged ' communicators ' include a group of deceased persons—Gurney, Sidgwick, and Myers, as well as Mary Lyttelton herself. Lord Balfour himself was present at some of the later sittings.

² Professor Ayer, I fancy, would now accept the general line of argument I have tried to illustrate here : in a more recent book, after giving " good reasons for supposing people do not survive their death ", he adds : " this is not to say that the notion of survival is self-contradictory or meaningless ; unless the hypothesis made sense one would not be entitled to say it was ' highly improbable ' ; for only what is possible can be false " (*The Problem of Knowledge*, p. 192).

by A's consciousness, and its absolute isolation from every other consciousness—peculiarities which are primarily matters of fact. This in itself is a sufficient reason, regardless of the apparent connection (an empirical fact, I take it) of A's consciousness with A's body. Indeed, it is in my view somewhat misleading to talk of sense-experiences in the plural, and then to inquire if and why they belong to one and the same self: such experiences are abstractions from what in fact is an indivisible continuum, which can even bridge the gaps we suppose to be caused by dreamless sleep.

Finally, so far as I can follow it, the whole doctrine appears to go round in a circle. 'Sense-contents' are said to be elements of a self. The identity of each self is defined by reference to a body, i.e., to a particular physical object. But physical objects are constructed by arranging sense-contents according to 'the laws of physics' while minds are constructed by arranging them according to 'the laws of psychology'. But if (as the context implies) mind here means consciousness, my consciousness is not a hypothetical constructional arrangement: it is a unique continuum, directly given not logically constructed; and in it what I call my sense-contents appear as elements fused or merged with other elements of a different type—mental images, feelings of pleasure or pain, beliefs, memories, wishes, and desires—which never as such enter into the construction of physical objects.

However, it is here superfluous for me to press my own objections to Ayer's earlier conclusions¹, since he himself (so he tells us) is now "doubtful whether the account given in [his] book was correct." In a later work on *The Foundations of Empirical Knowledge* he argues that "it is *not* logically inconceivable that I should have an experience that is in fact owned by someone else"; and from this he infers that the argument from analogy might after all be justified. But more recently still he appears to have developed doubts about even this further line of reasoning. As he puts it, he is now "inclined to revert to a behaviouristic interpretation of propositions about other people's experiences; but", he adds, "I own it has an air of paradox which prevents me from being wholly confident that it is true"².

Among British psychologists the writer who has most fully and frankly adopted the doctrines of logical positivism is Professor Mace. In his article on 'Physicalism'³ he undertakes to "defend an essentially physicalistic or behaviouristic view of the facts with which psychology deals". He begins with a brief criticism of 'radical dualism' as upheld, for instance, by Professor Spearman⁴. In place of this, he says, "I want one world of material things, and one science to deal with all that this world contains".

¹ I have ventured to mention them chiefly because so many students of psychology (and writers of textbooks as well) still accept the earlier views put forward in his influential book as lending authoritative support to behaviouristic doctrines.

² *Language, Truth, and Logic*, 1954, Introduction to new edition, p. 16.

³ *Proc. Aristot. Soc.*, XXXVII, 1937, pp. 23–40.

⁴ C. A. Spearman, *The Nature of Intelligence*, 1927, p. 66.

If I understand his description rightly, this one 'unified science' would consist in effect of what might be called a systematized chrono-geography of qualities. All our observations, he believes, can be reported in the form of what, following the positivists, he calls 'protocol statements'. Each and all of these may be symbolized by the general formula (Q, p, t), which translated into words, asserts that some specified quality (Q) pervades a specified place (p) at a specified time (t): e.g., 'cold here now'. Combined with certain supplementary statements ("visual, tactile, and other protocols") it will serve to verify the 'physical' proposition. "That is a block of ice." Combined with a different set of statements ("shivering here now", "aching here now") it will serve as a verification of the 'psychological' proposition, "I feel cold".

It is the aim of science, says Professor Mace, to derive laws. Thus, in physics, after making certain appropriate abstractions and establishing certain correlated variations, we arrive at formulae of the general type

$$C=f(T),$$

e.g., "when a bar of metal is heated, its length expands". Similarly in psychology it is our ultimate purpose, after carrying out the suitably planned experiments and subjecting the data to statistical analysis, to obtain laws, equally precise, expressed by analogous mathematical equations, e.g.,

$$S_1 - S_2 = f(I_1/I_2),$$

the well known 'law' which claims to state the way the intensities of two sensory qualities are functionally dependent on the intensities of two stimuli. "At this level", Professor Mace concludes, "no distinction can be drawn between psychology and physics: we have merely a neutral system of formulae." As for the 'protocol statements' which verify the laws, none of those used in what is ordinarily called psychology "requires for its verification any predicates of a type fundamentally different from those we employ in expressing our observations of material things"¹.

In all these proposals there is much that is bound to appeal to the statistical psychologist; and at the time they were published they provided a salutary check on the confused situation into which British psychology seemed to be drifting. More recently much the same arguments have been revived by Mr. Taylor and by other writers. Accordingly a few brief comments may be passed upon them here. The initial generalizations—that all branches of natural science, psychology quite as much as physics, are or should be based on first-hand empirical observations, and that these observations are in most cases of much the same formal type ("Here, now, such-and-such quality", or sometimes I should add,

¹ It may be noted that Carnap, at the end of the paper discussed above, adopts much the same standpoint. Having shown that "psychology becomes part of the unified science based on physics", he explains that by 'physics' we mean, not the system of currently known physical laws, but the science characterized by a mode of concept-formation which traces every concept back to *state-coordinates*, that is, to systematic assignments of numbers to space-time points". I fancy, much misunderstanding might have been avoided, had Carnap emphasized this somewhat unusual definition of 'physics' at the very outset of his defence of 'physicalism', and adhered to it throughout his argument.

"such-and-such a magnitude")—these assumptions appear to be beyond dispute. Nevertheless, I fail to see how they support the final conclusion, namely, that psychology at bottom is concerned solely with material things. After all, Fechner's law, as Professor Mace quotes it, expresses a relation between physical things and *sensory qualities*, whereas the laws of physics express relations between physical things and physical properties only.

But my chief trouble is to understand what exactly is meant by the words 'here' or 'the place *p*' as they occur in Mace's protocols. Is the 'place' to be determined by specifying the usual coordinates after the fashion of an aeroplane pilot—latitude, longitude, and altitude (as Carnap's statement might seem to imply)? Surely these can hardly appear in the basic protocols which record first-hand observations and nothing else; they are highly abstract constructions reached only at a relatively late reflective stage. Dr. Gallie, who seems to have been pulled up by the same difficulty, takes the word 'here' to mean "'occupying a place in this somatic sense-field', where 'this' is used as a proper name", and the 'somatic sense-field' is defined as a spatial volume which normally extends to "the surface of one's skin". Suppose then I am reporting my observations of a reddish star that I have just perceived about 10° above Orion's belt: is the protocol to be worded "Red here now"? Should we not rather say "Red *there* now"? In any case, I locate it far beyond the surface of my skin. Moreover readers of the protocol will surely want to know who made the report: was he a trustworthy reporter? Who or what is the subject¹ of the verb "observe"?

¹ I cannot here stop to discuss the objection that the wording of this question tacitly converts a convention of grammar into a postulate of science. I will merely say that the appeal to language can be used in two ways: we may argue that the traditional forms of our sentences are misleading; or we may argue that what has got embodied in almost every language expresses the most natural interpretation of everyday fact.

I myself should prefer to express both Mace's propositions in the same grammatical form, and report "I feel a block of ice" in the one case as well as "I feel cold" in the other. Admittedly the nominative pronoun would then suggest an actual subject and the verb a cognitive relation. Doubtless it was to avoid the appearance of thus tacitly implying a questionable epistemological doctrine simply on the basis of linguistic usage that Mace alters the natural form of speech. He himself does not stop to justify the change. But it was defended at great length by Russell. In his *Analysis of Mind* (1921) he rejects the idea of a cognizer and a cognitive relation (which he had formerly upheld) and argues that *sensa* just happen. This may seem plausible enough so long as we confine ourselves (as he commonly does) to cognitive experiences. But sensory experiences are not only sensed; they may be noted, remembered, wanted, like, or disliked. And what about the more complex experiences in which we catch ourselves believing, or reasoning about things, or reaching conclusions? Are we to say that our beliefs, and our arguments, and our conclusions simply occur, and they need imply no process of reasoning or believing and certainly no reasoner or believer? I fancy that many of the theoretical difficulties of philosophers and psychologists arise from concentrating almost exclusively on the cognitive aspects of consciousness—as though the sole activities of mind consisted in sensing sense-data, perceiving objects, inferring the existence of other minds by formal reasoning, and so on. The unity of experience is not a unity of cognitive contents, but primarily of purposive activities of which cognition is only one. The self is first and foremost a doer and only secondarily a knower. *Im Anfang war die That*.

One final piece of evidence for what it is worth. I myself am convinced that at times I actually cognize my cognizing, as well as my wanting, liking, my believing, etc. But, to avoid misunderstanding I hasten to add that the 'subject' with which I identify my observing 'self' need claim

The 'observer' might conceivably have been a spectroscopic camera in the nose-cone of a satellite transmitting its 'protocol' in code. But with the physicalist hypothesis in mind, let us suppose that it is my 'brain'. I see no reason for objecting, as Descartes would have done, that being a material thing extended in space is *logically incompatible* with being a mental thing able to perceive or think. But if a 'brain' is assumed to do the observing, then it will not be *merely* a material brain; it will not be just a convoluted lump of nervous tissue like that described in Watson's textbook on behaviouristic *Psychology*—the physical organ seen and stimulated by the surgeon in the operating theatre, dissected post mortem by the anatomist, and inspected under the microscope by the histologist: for we are now ascribing to it an additional property which material objects as such do not possess—the 'relational property' (or to borrow W. E. Johnson's term, the 'transitive adjective') denoted by the phrase 'being conscious of'. And, unless we tear the protocol right out of its context, the very existence of such a 'protocol' implies that the 'brain' had a number of other characteristics which the material things in general never exhibit, namely, the intention to watch and make an intelligible report—in short, conative and affective characteristics as well as cognitive¹. In that case the new qualities will be so distinctive that it is tempting to wonder whether we ought rather to speak, not of a brain endowed with mental characteristics, but rather of a mind which on the macroscopic level either manifests, or is in some way associated with, the characteristics of a material organ.

My final conclusion is therefore this. The sweeping generalization which the logical positivists have sought to uphold—that "all sentences of psychology describe physical occurrences"—proves to be quite untenable. No doubt, *some* sentences of psychology can be translated, without remainder, into the language of physics; but the most characteristic cannot, namely, conscious occurrences. "A man born blind", says Russell, "could know the whole of physics; but he could know nothing of colour. Hence the world of physics does not contain everything." Let me broaden this by saying that an 'electronic brain' could in theory formulate and apply all that we can know of physical occurrences; but it could never discover that some beings are actually *aware* of the occurrences that happen to them, unless one of those beings passed on the information.

none of the virginal purity of Kant's 'Pure Ego'. It might, for instance, consist merely of the concrete series of events that make up the one long event I call my mental life: in other words, if the events or occurrences were the *processes* rather than the *sensa*, then they themselves could serve as 'subject'. And half a dozen other hypotheses are possible without postulating some transcendental noumenon.

¹ Here let me note that Professor Ayer at any rate does not repudiate the idea of a 'cognitive relation' as I have called it, since he explicitly refers to the "relation of being experienced by a particular subject"; and Professor Mace emphasizes the importance of 'conative tendencies' in his discussion of belief, and comments on the fact that such tendencies "involve difficulties for simple behaviouristic theories".

If we are to accept the existence of the physical world, as ordinarily conceived by the physical sciences, this conclusion would manifestly commit us to the dualism which contemporary philosophers and psychologists are so anxious to avoid. But surely the whole history, both of physics and philosophy, has demonstrated again and again that the universe does not possess that simple homogeneous nature for which the passionate monist is for ever pining. And after all we can assert a reasonable degree of unity without denying an equally fundamental variety¹.

The concluding paper in the book brings the series to an unexpected climax. The story opened with Bertrand Russell's claim, set forth on the very first page, that "logic is what is fundamental in philosophy"; Waismann ends it by declaring that "the fact that a philosopher's arguments may not be logical does nothing to detract from their rationality: he may see an important truth, and yet be unable to demonstrate it by formal proof". This is a strange pronouncement in the mouth of a former logical positivist. *Nam et Saulus inter prophetas?* Nevertheless, the psychologist, who is only too well aware of the gaps in his own demonstrations, should be the first to welcome it. Some of Dr. Waismann's pregnant remarks might well be taken to heart by the more stringent methodologists who have of late attempted to instruct the psychologist how he should plan his experiments and formulate his proofs. What is wrong, (we are told) with the way discoveries are commonly presented is that they are set in a slick and systematic framework as though they were the inevitable outcome of some logical 'method' or 'procedure'. "Great men seldom arrived at their solutions by drawing rational inferences: there is nothing like clear thinking to prevent one from making discoveries." The work of a scientist in his laboratory ought, he suggests, to be modelled rather on that of a court of justice than on that of someone "doing sums"; and "there are no computing machines to perform the functions of a judge". It is an aphorism which every statistical psychologist might well hang over his desk.

¹ If we give up the notion of an independent physical world instead of accepting it as basic (as the orthodox positivists would have us do), the dualism still remains. Even Russell—the most lucid and logical exponent of 'neutral monism'—cannot completely dispense with it: the only difference is that it now becomes "a dualism, not of entities, but of causal laws . . . One kind, belonging to physics, the other to psychology" (*Analysis of Mind*, pp. 25–6, cf. p. 137).

NOTES AND CORRESPONDENCE

THE FACTOR ANALYSIS OF THE WECHSLER SCALE. I

The Evidence for a Simple Structure

By M. A. JACKSON

In a recent issue of the *British Journal of Educational Psychology* (XXIX, pp. 237f.) Dr. A. E. Maxwell has criticized the customary method of classifying the tests included in the Wechsler Intelligence Scale for Children. "If" (he very rightly observes) "it is thought desirable to divide the tests into subgroups, each measuring a separate aspect of cognitive ability, it would appear preferable to proceed along factor-analytic lines: in this way it becomes possible to discover just how many independent factors are necessary, and, after suitable rotation, to label them in a meaningful way." He has accordingly factorized the three correlation matrices published by Wechsler himself, each based on the performances of 200 American children; and he claims to have shown that the figures can be interpreted in terms of two group factors only, representing "relatively pure aspects of cognitive ability", namely, a 'verbal intellectual' factor and a 'space performance' factor.

In point of fact, however, Wechsler himself has already published the results of several factorizations of his scales and given results both for children and for adults¹. Of the various investigators, some (he points out) have adopted Burt's orthogonal procedure, others Thurstone's centroid method, using either oblique or orthogonal rotation. But with all the investigators and for each of the scales much the same factorial patterns are obtained. So far from obtaining only two factors, Cohen obtained as many as six. The factors most frequently reported are a general factor, usually identified with intelligence, and two broad group factors for verbal comprehension and for visuo-spatial performance respectively. Others factors include memory, numerical facility, conception or reasoning, mechanical ability, and a factor for speed and concentration ('capacity to resist distraction', reported by Cohen²).

Dr. Maxwell begins by factorizing each correlation table according to the method of maximum likelihood, and, having described and tabulated the relevant results, he states that "in this way it was possible, after suitable rotation to label them [the factors thus obtained] in a meaningful way". The reader might take this to imply that, after a series of trials, the mode of rotation ultimately selected was the only "suitable" rotation that would yield a meaningful set of factors. Actually, however, his procedure would seem to have been quite different, and indeed somewhat unusual. "The first factor in each analysis," he explains, "was made to pass through Vocabulary; the second factor was then kept orthogonal to it." This apparently arbitrary decision is explained by saying that it was carried out "for reasons of expedience". But what precisely the reasons were we are never told. Naturally the first of the factors so extracted has a high loading for the vocabulary test; and, in discussing the figures obtained, Dr. Maxwell argues that "this test gives us an indication of the child's knowledge of words, and therefore provides a lead for interpreting the factor". But it seems evident that he must have approached

¹ D. Wechsler, *The Measurement and Appraisal of Adult Intelligence*, 1958, chap. VIII, 'Factorial Composition'. So far as I am aware only one investigator has factorized the scale for children (U. Gault, *Aust. J. Psychol.*, VI, 1954, pp. 85-90). But, as Wechsler's tables show, the factorization of the adult scales yields much the same results.

² J. Cohen, *J. Abnorm. Soc. Psychol.*, XLVII, 1952, pp. 359-365 and *Amer. Psychol.*, XI, 1956, pp. 449 f.

the data at the outset with a preconceived theory about the nature of the main factor underlying the scale, namely, that it would depend chiefly on the child's mere knowledge of words¹; and he therefore deliberately selected a position for the factorial axis which would yield the 'meaning' he desires. But plainly such a rotation can prove nothing.

A study of Wechsler's original tables will show that, in all three age groups, the test which accounts for the largest amount of correlation is not the Vocabulary test, but the Information test; and several critics have consequently complained that the scale depends far too much on acquired information. A factorist with this alternative hypothesis could, by adopting Dr. Maxwell's principles, and 'passing the first factor through' the Information test, demonstrate his theory quite as easily as Dr. Maxwell has demonstrated his.

So far from being 'pure', the two factors Dr. Maxwell has obtained seem manifestly composite. This is apparently recognized by Dr. Maxwell himself in the earlier part of his paper. His first factor, he tells us, is not 'simply *verbal*' (his italics). And, since it enters into every test, it must include what is virtually a 'general factor'. Hence presumably its composite name. Furthermore, since his chief aim was to ascertain the number of factors required, we should expect him at least to examine the residuals left after his factors have been extracted, and to test their statistical significance. The sample of children tested by Wechsler was unusually large, 600 in all; and, on comparing the products of Dr. Maxwell's factor loadings with the original correlations, appreciable residuals remain over, several of which appear to be fully significant.

However, what English users of the Wechsler scale require to know is, not what kind of abilities are used by American children in solving the tests, but what kind of abilities are measured when the tests are applied to English children, or at least whether it is safe to assume that the same factors underlie the performances of both. We believe that some light may be thrown on this problem by some of the data we ourselves have collected for English pupils. We have here confined ourselves to the first six tests in the Wechsler scale, and so far have obtained data from approximately 200 children aged 10 to 11 (86 boys and 121 girls)². There were several reasons for excluding the so-called 'performance' tests from this part of our study. First of all, 'performance' tests, and the abilities underlying them, are of comparatively little interest to the ordinary teacher. Secondly, the inclusion of such tests in the same factorial battery tends to obscure some of the more specific factors: the 'general factor', for instance, becomes a composite of both verbal and non-verbal abilities, and probably of other abilities as well, all of which ought to be kept separate. Thirdly, as has been pointed out by Burt and others, in a battery of tests intended for practical use in a given type of situation, what is of importance is, not so much the 'absolute value' of the tests selected, but what he calls their 'relative value'. "Battery A+B may have a higher theoretical value than Battery A alone, when judged in the usual way by a so-called 'validity coefficient'; and yet Battery A by itself may be preferable (i) because for the purposes in question it adds just as much, and possibly more, *to the information already available*, and (ii) because, being shorter, it is more economical of time, trouble, and expense."

Our aim in factorizing the data was partly the same as Dr. Maxwell's, namely, to ascertain the number of independent factors present. But we also hoped to obtain by

¹ Mr. Maxwell omits two of the twelve tests (Information and Mazes), but, oddly enough, adds the total scores for the 'verbal' and the 'performance' tests as though each total score formed a single test. This may in part account for the discrepancy between his results and those of other investigators.

² The boys were tested by Mr. K. J. Roberts and the girls by Miss Bannister and myself. When further data have been collected we propose to apply the same methods to the *complete* battery; but we publish this preliminary note to begin with, in the hope of eliciting comments and criticisms from other workers who may have used (and possibly factorized) the Wechsler scales.

means of a 'suitable rotation' a set of 'primary abilities' in Thurstone's sense, i.e., what he terms a 'simple structure' consisting of broad and narrow 'group factors'. We began by making a *complete* 'centroid analysis', i.e., by continuing the extraction of factors until all the residuals were reduced to zero. We followed the standard procedure¹, entering in the diagonal the largest coefficient (or residual) in each column and 'reflecting' signs in the usual way. With this procedure six factors were needed to complete the analysis. The variance contributed by the last three is comparatively small; but they are helpful in securing the best possible fit for the rotated factors. The rotation itself was carried out according to the method described by Thurstone (*Multiple Factor Analysis*, chap. X).

TABLE I. CORRELATIONS

Tests	Tests					
	Inf.	Comp.	Arith.	Sim.	Voc.	Dig.
Information	—	.661	.440	.547	.769	.477
Comprehension	.661	—	.547	.744	.584	.455
Arithmetic	.440	.547	—	.429	.400	.612
Similarities	.547	.744	.429	—	.480	.335
Vocabulary	.769	.584	.400	.480	—	.437
Digit Span	.477	.455	.612	.335	.437	—

TABLE II. ROTATED FACTORS

Tests	Factors					
	V	R	M	N	W	G
Information	.683	.102	.497	.184	.248	.131
Comprehension	.635	.539	.153	.149	.112	.323
Arithmetic	.164	.508	.464	.208	— .124	.157
Similarities	.558	.481	.086	— .007	.104	.366
Vocabulary	.507	.097	.545	.114	.405	.194
Digit Span	.156	.511	.596	.243	— .036	— .132

The initial correlations are shown in Table I, and the rotated factors in Table II. The loadings which appear statistically significant are printed in heavy type². The first factor (V) is plainly a factor for verbal relations; the second (R) appears to be a factor for Reasoning; the third (M), with its high loadings for Vocabulary, Digit Span, and Information seems to correspond with Cohen's Memory factor; of the narrower factors the one marked N, with high loadings for Arithmetic and Digits is obviously a Numerical factor; the next, with its highest loading in the Vocabulary test, is a factor for Words (Thurstone's W as distinguished from V); and the last (G) might perhaps be regarded as a factor for Generalization.

¹ L. L. Thurstone, *The Vectors of Mind*, pp. 108f. For arguments in favour of a complete factor analysis see Warburton, this *Journal*, VII, pp. 101-106.

² In judging the significance of the factors we have applied the usual test to the residuals on which each factor is based. Thus (to take the most doubtful of all) after extracting the first five factors we find there is a residual correlation of $.753 - .636 = .117$. With 207 persons in the sample tested the standard error of an observed correlation of .753 would be $\pm .030$. Judged by this criterion the residual appears fully significant. We are aware that this method of testing the significance of a factor is opening to criticism; but so are nearly all the other methods. Moreover, it may seem a little rash to extract six factors from only six tests; but we consider that, in an exploratory study, it is of greater interest to discover how *many* different abilities enter into the scale, not how *few*. The demonstration of their full statistical significance must await the collection of a larger mass of data.

There is nothing very novel about the factors thus discovered. The chief value of the investigation arises in our opinion from two main features : first, we have applied factorial methods to data obtained from British children with six of the Wechsler tests and shown that, so far as they go, the main results (except possibly for the last rather doubtful factor) tally fairly well with the results obtained from American children ; secondly, we have shown that the correlations can be interpreted in terms of a ' simple structure ' of ' primary abilities ' *without invoking any general factor*.

THE FACTOR ANALYSIS OF THE WECHSLER SCALE. II

By CYRIL BURT

The comments which Miss Jackson has made on Dr. Maxwell's analysis of the Wechsler Intelligence Scale for children are highly instructive ; and most teachers, I think, would agree with her criticisms of his conclusions. On the other hand, her own factorial procedure seems open to equally serious objections ; and I myself am tempted to say that her tables provide unusually clear illustrations of what many of us have always felt to be the defects or dangers of Thurstone's methods¹.

First of all, however, may I point out that the phrase ' complete analysis ' is somewhat ambiguous? Dr. Warburton and others have recently argued in favour of the procedure whereby as many factors are extracted as there are tests (this *Journal*, VII, pp. 101f.) ; and this is not infrequently referred to as a ' complete analysis '. But their procedure commonly begins by inserting units in the leading diagonal. Miss Jackson, however, inserts not units but the largest coefficient in the column. To avoid confusion, therefore, Dr. Warburton's method might perhaps be more conveniently termed a ' full factor analysis '—the phrase Dr. Warburton himself usually employs. The insertion of units implies that no specific factors are present.

When, as in the case of the Wechsler tests, specific factors peculiar to each component of the battery are almost certainly operative, then ' reduced self-correlations ' must be substituted ; and the size of these will depend on the variance of the common factors. In this case a ' complete analysis ' can only mean a complete analysis of the *observed* correlations, i.e., of the non-diagonal coefficients ; and hence the largest number of common factors we can legitimately assume will be equal, not to the total number of tests, but to the minimum rank permitted by the observed correlations. When (as here) the number of tests, n say, is a ' triangular number ', i.e., when $n = \frac{1}{2}m(m+1)$ where m may be any integer, then the minimum number of factors required for a ' complete analysis ' will be $r = n - m = \frac{1}{2}m(m-1)$, and the self-correlations (or ' communalities ') will be uniquely determined. Thus, so far as the statistical data contained in the correlation table are concerned we have no grounds whatever for postulating more than r factors ; and of course, if the number of persons tested is strictly limited, some of these may be statistically non-significant. There may perhaps be extraneous reasons for postulating a larger number ; but, if so, the reasons should be convincingly stated. It follows therefore that, apart from such extraneous evidence, the maximum number of factors we can assume for a battery of six tests only will be three².

¹ See *Factors of the Mind*, 1940, pp. 361f. and 450 ; also this *Journal*, IV, pp. 145 and refs., and V, pp. 168f. I am much indebted to Dr. Maxwell and Mrs. Sheila Jones for their comments on Miss Jackson's paper. In the main, I gather, they would be inclined to endorse my criticisms of her procedure.

² Cf. *Factors of the Mind*, 1940, pp. 109, 460-1. In principle, it would seem, Thurstone accepted the argument set forth above, since in a footnote in his later volume he observes : " operationally, it is correct to say that, when the communalities have been written in the diagonal cells, the rank is a minimum . . . The communalities maintain the rank of the side correlations, and this is the minimum rank " (*Multiple Factor Analysis*, 1947, p. 283).

To achieve a perfect fit with no more than three factors, the *exact* figures for the self-correlations must be inserted in the diagonal when calculating the loadings for the first factor; (the values here required are shown in the first column of Table I A). And, when calculating values for the later factors, the residuals in the diagonal must be left unchanged. Thurstone's method of inserting the largest coefficient in each column will *not* produce the best possible fit; after three factors have been extracted further residuals are bound to remain, and it is pointless to test these residuals for statistical significance, since they are mere artefacts—errors of approximation resulting from the arbitrary way in which the self-correlations have been estimated. The proper working procedure is to recalculate the first three factors substituting better estimates for the self-correlations (namely, the square-sums of the factor loadings thus obtained) in the leading diagonal, and to keep repeating the whole process until the exact values are secured. Moreover, if, in accordance with Thurstone's procedure, the largest residuals are inserted at *every* stage of the analysis, then the errors will, as a rule, be cumulative; and in the end we may easily obtain spurious residuals which are more than twice the standard error of the observed correlations¹ and thus appear fully significant. This is one of the chief reasons why Thurstone and his followers so often obtain far more factors from their tables than other investigators have done. In *Primary Mental Abilities*, for example, he tabulates, and seeks to interpret, as many as 13 rotated factors for a set of tests applied to only 218 students².

To demonstrate my point I have in Table I shown how, with appropriate communalities, the table of correlations can be exactly fitted with three factors only. The reader who cares to carry out a check will find that (except for tiny discrepancies of about .001 or less, resulting from rounding) all the residuals are precisely zero. Such a fit can only be achieved with three factors when two of them are 'bipolar', i.e. include negative loadings which ordinarily will be of about the same number and size as the positive.

However, in dealing with cognitive abilities it is more natural to think of the factors used to represent them as exerting a favourable effect on the processes into which they

TABLE I. FACTOR SATURATIONS

A. Bipolar Factor Matrix					B. Group Factor Matrix			
		Factors				Factors		
Tests	Comms.	I	II	III	I	II	III	IV
Information	.854	.835	-.240	.316	.785	-.006	.488	.015
Comprehension	.838	.853	-.168	-.288	.818	.038	.027	.410
Arithmetic	.668	.689	.426	-.106	.610	.533	-.077	.073
Similarities	.675	.715	-.231	-.332	.698	-.053	-.012	.433
Vocabulary	.693	.749	-.198	.304	.702	.011	.447	-.006
Digit Span	.603	.650	.411	.106	.568	.519	.074	-.076

¹ This by itself is a far from safe test for the significance of factor loadings; cf. C. Burt, 'Tests of Significance in Factor Analysis', this *Journal*, V, pp. 125f. If the test there proposed on p. 126 is applied, very different conclusions will be reached.

² It is only right to add that Thurstone and most American factorists commonly work with much larger batteries than British investigators; and this to some extent minimizes the effects of any crude approximation. Moreover, in his later volume (*op. cit.*, chap. XIII, p. 300) Thurstone does in fact show that he appreciates the defects of this "simple method of estimating communalities", and indicates the need for improved estimations when the batteries are small. But the point is not brought out in his routine description of the procedure in chap. VIII; and many research students seem not to read as far as the later chapter. In practice Thurstone and those who followed his methods seldom attempt to extract the full number of factors which their own residuals would permit; and hence the errors introduced by their rough and ready treatment of the diagonal values are apt to escape notice. Here the adoption of his method with an unusually small table of correlations brings its defects into clear relief.

enter, and not a negative or inhibitive effect ; and this implies that all the significant loadings should be positive. How then are the negative loadings to be eliminated? The method I proposed was based on the following principles. Factors are essentially principles of classification ; and the pattern of positive and negative signs in the bipolar factor matrix indicates the way in which the successive classes are subdivided. Thus, when (as usually happens with tests of cognitive abilities) all the observed correlations are positive, the first factor necessarily has positive loadings throughout. It thus represents the general class or *genus* to which all the tests in the battery belong. To determine its precise nature we should really need other tests in the battery, e.g., performance tests and some well-authenticated test of intelligence, such as the revised Stanford-Binet : as it is, we may hazard a guess that the general factor here obtained represents a mixed ability for intelligent comprehension and acquired information ; let us call it an 'intellectual' factor. The second factor is bipolar : it divides the entire group into two main classes or species. It has its highest loadings in the two numerical tests, and appears to distinguish these from the non-numerical. To express this classification in positive terms, we need two group factors ; the first will have (if we can manage it) zeros wherever the four negative loadings appear ; the second will have zeros where the two positive loadings appear, while the negative loadings will have to be converted into positive. The third factor subdivides this latter group into two subspecies, each of which will have to be represented by a separate group-factor. And so we can continue until we reach the most specific of all—the so-called 'specific' or 'individual' factors each entering into only a single test.

These further conditions will clearly entail a larger number of factors than were contained in the 'bipolar' matrix (Table III). Ignoring factors that are wholly specific, and using the phrase 'group factor' to include all factors common to two or more tests (possibly all), it is evident that the largest number of group factors we may need to provide a complete representation of a given matrix of bipolar factors will be $k = (2^r - 1)/(2 - 1)$, where r denotes the number of factors in the bipolar matrix. It might be supposed that this would require $n \times k$ observed coefficients. That, however, is incorrect ; the conditions just laid down fix in advance the values for many of the $n \times k$ elements in the new factor matrix, since those conditions require that certain of the elements in certain specified positions shall have the value zero (or approximately zero) and limit others.

Thus a complete group factor matrix corresponding to the scheme of classification suggested by the bipolar matrix set out in Table IA would embrace the following factors : (i) a general or 'basic' factor, subdivided into (ii) a broad group factor covering the four verbal tests, and (iii) a narrow group factor covering the two numerical, the former being subdivided into (iv-v) two narrower group factors, each containing two tests only. However, a glance at Table IA indicates that the broad group factor would have very small loadings of doubtful significance. Let us therefore provisionally omit this, and calculate the other factors first of all, and then examine the residuals.

Accordingly we begin by making a fresh analysis of the table of correlations into a set of non-overlapping group factors, dividing the groups according to the scheme indicated by the preliminary bipolar (or 'centroid') analysis¹. This yields a single basic factor and

¹ For the working method see *Factors of the Mind*, 1940, pp. 477f., or this *Journal*, III, pp. 45-49. In calculating the new 'rotated' factors, it is advisable to adhere to the principle that guided the calculation of the original factor matrix, i.e., we shall calculate first of all those factors that make the largest contribution to the available variance, and, if any of the k factors are to be omitted, then we shall omit those that would make the smallest and least significant contributions. By using a linear transformation we can always convert one set of uncorrelated factors into another set (embodying certain acceptable specified conditions if desired) from which we can still reproduce the initial correlation matrix. Such a transformation will be orthogonal, i.e., it will (to use the familiar term) constitute a 'rotation'. The idea of rotating factors was due originally to Maxwell Garnett ('On Certain Independent Factors in Mental Measurement', *Proc. Roy. Soc., A*, XCVI, 1919, pp. 91-111 ; cf. *id. Brit. J. Psychol.*, IX, 1919, pp. 345-366).

three non-overlapping group factors. After these have been extracted, the residuals prove to be much too small and irregular to justify any further factors, though they suggest some small degree of occasional overlapping. A formal rotation therefore seems desirable.

To avoid introducing subjective preconceptions into the process of rotation it is essential that this rotation should be carried out automatically by means of an arithmetical calculation. A quick and easy method is described in the publications just cited¹. From the matrix of non-overlapping factors a hypothetical table of correlations will already have been reconstructed in order to check the closeness of fit. If, as we have just seen, the residuals are too small to warrant further factorization, this hypothetical table must be factorized by the summation (or 'centroid') method used for the observed correlations. It is now a simple matter to compute the orthogonal rotation matrix required to rotate this new hypothetical centroid matrix into the hypothetical matrix of group factors (cf. *loc. cit. inf.*, Appendix II). The 'rotation' matrix thus obtained is shown in Table II. This transformation matrix is applied to the original centroid matrix (based on the observed correlations, and shown in Table Ia) ; and in this way we reach the final matrix of overlapping group factors shown in Table Ib.

TABLE II. ROTATION MATRIX

·9305	·2320	·2110	·1893
—·0814	·8854	—·3404	—·3054
—·0362	·0409	·7280	—·6834

The factors thus obtained may be interpreted as follows. There is first a basic 'intellectual' factor, common to all the tests. This takes the place of Miss Jackson's three broad overlapping factors. Superimposed on it are three narrower factors corresponding roughly to the last three factors found by Miss Jackson, but far larger in size. They show less overlapping than Miss Jackson's factors ; and in this sense appear to represent 'purer' cognitive abilities (if I may borrow a phrase from Dr. Maxwell). Of these the largest is the 'numerical' factor (with Miss Jackson's analysis it was the smallest of all) : it overlaps slightly with Digit Span, but the figure is scarcely significant. The next largest is a 'verbal' factor (corresponding to Thurstone's W). The smallest seems to be a factor for 'Comprehension' : this shows a slight overlapping with Arithmetic, but again the figure is scarcely significant. From these four factors the observed correlations can be reconstructed exactly : there are no residuals.

The advantage of this procedure is that at every stage the working is completely automatic and objective. The computer's idea of what is or is not meaningful in no way affect his results. If in the end those results agree with his ideas, then he may safely feel he has objective evidence to support them, and that any other computer, who might have different ideas about what is meaningful, would arrive at the same results as he has done. The disadvantage of the method is that it requires an ability to solve linear equations ; and, as described in words, sounds rather laborious. A little practical experience, however, will quickly show that these drawbacks are not very serious².

¹ This *Journal*, III, pp. 61-71, and Appendix II.

² It is frequently objected that with large batteries of tests the recalculation and analysis of the hypothetical correlation is a lengthy task. But in such cases a summed matrix should be substituted for the detailed correlation matrix. Here, for example, since no group factor contains less than two tests we can compute the rotation matrix from a 3×3 table of summer correlations : (cf., this *Journal*, III, pp. 64f ; here I should call the reader's attention to the copying error in Table XVIIb on p. 67, corrected in the *Erratum* on p. 128 of the same volume). Let me add that, although considerations of meaning should not influence the factor analysis once the data have been collected, they should influence the initial planning of the investigation.

Thurstone's method of treating the problem, admirably illustrated by Miss Jackson's work, is widely different, and to the beginner undoubtedly seems far more attractive. Thurstone holds that the factors obtained by summational methods, whether obtained by 'simple summation' (as in the 'centroid' method) or by 'weighted summation' (as in the method of 'principal components', of which the 'method of maximum likelihood' is a special form) "are not psychologically meaningful". He therefore makes no use of the classification indicated by the pattern of positive and negative signs. Instead, in order to secure a unique rotation, he suggests that the matrix of rotated factors should conform, not to the scheme revealed by the bipolar matrix obtained from the particular correlation table that is being factorized, but to a general scheme, termed a 'simple structure', which is the same for every set of correlations. Its essential characteristics are, first there must be one or more zeros in every factor (if possible r zeros in each), i.e., there can be no general factor; secondly, each test must also have a zero loading for at least one factor.

Now, even in their most stringent form the conditions which define a 'simple structure' do not (as Thurstone originally supposed) suffice to make the rotation unique. As often has been demonstrated, with one and the same correlation table a number of different 'simple structures' can often be secured, since each investigator is free, within the limits imposed by the data, to place his zeros more or less as he wishes, and is consequently able (particularly when 'oblique' factors are permitted) to choose his factors to fit whatever 'meaning' he prefers. Here too it is by no means difficult to devise other factor matrices obeying the conditions of 'simple structure' which will account for the observed correlations quite as well as that proposed by Miss Jackson. What then are the reasons for preferring the matrix shown in Table V? First, the number of factors required is far smaller. Secondly the classificatory scheme which these factors imply has been automatically deduced from the scheme indicated by the bipolar analysis. Finally, we can in this case argue that the results agree with those of other analyses not only of the Wechsler tests, but also of cognitive tests generally.

The chief merit that Miss Jackson can claim for her scheme is that it conforms to the requirements of a 'simple structure'. But is this a genuine advantage? Why should we expect any and every battery of tests to exhibit this particular pattern to the exclusion of any general factor? If we could reasonably revive the old phrenological notion of a brain operating through more or less independent 'organs' or 'faculties', such an assumption might perhaps seem plausible. But all we know of the brain suggests, so far as cognitive processes are concerned, the joint operation of *both* a general and specialized factors.

However, quite apart from our views on the existence or non-existence of 'general intelligence', it is easy to show that the persistent demand¹ for 'simple structure' is from its very nature impracticable and self-contradictory. Several earlier factorists have analyzed one or other of Wechsler's correlation tables along Thurstonian lines, and thus obtained a rotated matrix showing 'simple structure'. Nearly all the 'structures' include two broad and slightly overlapping group-factors, first an 'intellectual', 'educational' or 'verbal' factor, covering the first six tests, and secondly a 'practical', 'performance' or 'space' factor, covering the remainder. Indeed, Dr. Maxwell himself seems to have had a structure of this kind in view when making his analysis. Now, had they followed Miss Jackson's plan of factorizing the first six tests only, then the 'intellectual' or 'verbal' factor which originally covered only a *group* of tests ought, if it is invariant, to appear as a *general* factor; but that of course would conflict with the Procrustean demand for simple structure. And suppose Miss Jackson were to carry her principle of exclusion still further and omit Arithmetic and Digit Span, then her own first group factor would become a

¹ As I understand his writings, this invariable demand is rather a requirement imposed by Thurstone's followers than a rule laid down by Thurstone himself.

general factor, and the principle of 'simple structure' would again break down. Clearly whether or not a 'simple structure' is to be expected must depend on the initial choice of tests. Wechsler's intention in constructing his 'Scales' was to compile a battery primarily to measure the "one pervasive factor in all performances requiring intellectual ability": why then should factorists go out of their way to eliminate this general factor, and substitute three or more group factors instead?

Miss Jackson is quite right in arguing that it is a little rash to assume, without more direct empirical evidence, that a factor-pattern obtained by factorizing data from American children will also hold good of results obtained with English children. But Wechsler's scales have already been applied to children in this country by Miss Mildred Marshak and Mrs. Sheila Jones. Both have factorized correlation tables obtained from various age-groups; and both have found in every case a general factor of 'intelligence' entering into the whole set of tests as well as two broad factors—an 'intellectual' factor and a 'practical'. However, with their data the line of division between the two by no means invariably agrees with Wechsler's classification: not infrequently Arithmetic and Digit Span tend to combine with Picture Arrangement and Coding—a classification which throws doubts on Miss Jackson's selection of the first six tests only. In addition, particularly with the older children, there are indications of smaller group factors, such as those disclosed here for numerical ability, verbal ability, and the like; but, owing to the small size of the groups, the evidence for these was non-significant¹. To pool the various batches of children is scarcely practicable since the pattern of the correlations varies somewhat with the different groups. The differences apparently depend partly on the age and sex of the children, but still more on the variations in education given in different British schools, on the differences in the type of problem at different age-levels, and on the differing reliability of the tests at different levels. With the younger children, the general factor tends to dominate over the more specialized factors; with the older the latter are more conspicuous, but with the Wechsler scales they are considerably influenced by the home and school background². A striking instance is the way the arithmetic test (when given to older children) often gets grouped, not with the other numerical tests, but with the verbal, owing partly to the verbal nature of the last two questions. Plainly, therefore, it is quite unjustifiable to assume that one and the same factor pattern will appear at all levels and with all types of children.

THE GENERAL AESTHETIC FACTOR. I

To the Editor, British Journal of Statistical Psychology

SIR,—In several papers contributed to this and other journals it is apparently assumed that a specific mental ability called the 'general aesthetic factor' has been definitely established: it is even claimed that those individuals who have high factor measurements for this particular ability are "usually of an intuitive type" (this *Journal*, XI, p. 145). However, the empirical evidence for such views seems far from convincing. The nearest approach to an objective proof is that contained in a paper by E. D. Williams and others published in *Brit. J. Educ. Psychol.*, VIII, p. 282. But their arguments are based not on a

¹ M. D. Marshak, 'A Study of the Wechsler-Bellevue Intelligence Scale as applied to British School Children', 1955, and S. Jones, 'A Study of the Wechsler Intelligence Scale for Children', 1959: Ph.D. Theses, University of London Library.

² Consider not only the problems in the arithmetic test, but also such questions as 'Who wrote *Romeo and Juliet*?' 'What is a hieroglyphic?' (Information) 'Why should most government positions be filled through examinations?' (Comprehension) 'What is meant by—"Aseptic?" 'Nitroglycerine?' 'Stanza?' 'Vesper?' (Vocabulary). These are hardly tests of 'intelligence' as the term is commonly used by British psychologists, and must vary greatly with the schooling and home background of the child.

formal factorial analysis, but merely on partial correlations. They add that their own figures "largely confirm Burt's inference that in the vast majority of persons aesthetic appreciation is dependent on a 'group factor' common to all the various media". Burt's discussion of the matter in the volume which they cite (*How the Mind Works*, 1933, chapter on 'The Psychology of Art', pp. 266f.) is mainly theoretical. "The sense of beauty", he writes, "depends essentially on the ability to perceive objects or sensations—shapes, colours, sounds, and even emotions—in certain relations" (*op. cit.*, p. 302, his italics). But have either he or Miss Williams and her coworkers obtained any actual factorial data to verify this rather speculative hypothesis? Surely to establish a *factor* we need something more than mere correlations; we need an analysis into factors.

In a later research, also carried out in Professor Burt's laboratory, Dr. Eysenck claims to have established a factor "more general than Dewar's but less extensive than that of Williams, Winter and Woods . . . extending over other modalities as well". But this can hardly be regarded as a *general* aesthetic factor. All his aesthetic tests were visual; the only other 'modality' tested was smell. Nor does he print the correlations on which his conclusion was based: they were, it would seem, simply the raw coefficients, factorized without any attempt to eliminate the effects of intelligence ('The General Factor in Aesthetic Judgments', *Brit. J. Psychol.*, XXXI, pp. 94–102). Several later investigators, using factorial methods, have reported the presence of a general factor common to musical, poetic, and pictorial appreciation; but they do not appear to have demonstrated that this factor is anything more than the ubiquitous factor of general intelligence. Moreover, even if it be maintained that there is indeed a further factor distinguishable from intelligence, what evidence is there to show that it is not after all an environmental artefact—the mere effect of cultural conditions—rather than a special type of innate ability, as seems so widely assumed?

The whole question is plainly of great practical importance to those of us who are engaged in educational or vocational guidance, and have to advise and possibly test eager aspirants for an artistic, musical, or literary career.

C. R. PATERSON.

THE GENERAL AESTHETIC FACTOR. II. (Reply)

By E. D. WILLIAMS

Mr. Paterson has undoubtedly put his finger on the chief difficulty confronting those who support the hypothesis of a general aesthetic factor, namely, the need to demonstrate that it is not just another manifestation of the "ubiquitous factor of general intelligence". The joint article of ours that he quotes dealt with his particular problem quite briefly, because (as its title indicates) it was concerned primarily with 'literary appreciation'. In the course of the investigations there described, Miss L. R. Bennett and Miss M. Wright tested a large number of children, and a smaller number of students, with other forms of aesthetic appreciation. The results were to have formed the subject of a separate paper; but, as was stated in the article, Miss Wright was recalled to Canada before this part of the research was completed; hence the data actually obtained were much more meagre than had originally been planned.

The tests for the three main subjects (literature, painting, and music) were, it may be remembered, of three main types: (i) ranking specimens in order of preference; (ii) paired comparison with specimens of somewhat similar types; (iii) choosing the preferred specimen from three or more versions of the same drawing, tune, poem, or prose passage. Tests of intelligence, both verbal and non-verbal, were also applied, and quasi-partial correlations computed by eliminating the effects of intelligence in accordance with the formula $r_{ab.g} = r_{ab} - r_{ag}r_{bg}$ (where g denotes general intelligence). With the more usual formula for partial correlation this residual coefficient $r_{a.g.b}$ is divided by the product of the 'alienation coefficients'—a procedure which considerably magnifies its apparent value

and is even liable to produce spurious factors as a result of the correction. We therefore factorized the residual coefficients without correction. The method used was the method of 'group factors', with the groups determined by a preliminary bipolar analysis. At that time, however, all these procedures were comparatively new, and had been somewhat severely criticized by various writers¹. Partly for this reason, therefore, and partly to save space, these further tables were omitted in the paper to which Mr. Paterson refers. Instead we calculated the correlations between the various factor measurements, and then partialled out the effects of intelligence, at the final stage instead of at the outset, by means of the more familiar formula, which our critics regarded as less objectionable.

Today, however, the factorization of residual coefficients is a recognized practice, and the method of group factor analysis, based on rotation, is a well-established technique. Accordingly, we give below the detailed correlations and the factorial results. Table IA gives the figures so obtained in our own investigation; and for comparison we reprint in Table IB the results obtained by Burt and Pelling from a group of 168 elementary school-children aged 11 to 12. The conclusion which we quoted from Burt was, as was expressly stated in our paper, his "inference from *his earlier test-results*", not the somewhat speculative inference put forward in his book, which was meant solely for the general reader.

The earlier investigations were carried out, not only on children of school age, but also on young adults attending schools of art and evening classes or applying to the Vocational Section of the National Institute of Industrial Psychology. With the latter the numbers tested were considerably larger and tests included a wider range of subjects².

In all three groups a general factor for aesthetic appreciation was clearly demonstrated; but it appeared to play a more conspicuous part in the older group. This perhaps is partly the result of maturation during adolescence, and partly no doubt as a result of cultural conditions operating during later years. Burt indeed in one of his earlier publications inquired whether a behaviouristic critic might not argue, with some plausibility, that the whole of the general factor resulted from cultural conditions.

It was to answer this latter problem, rather than to demonstrate afresh the existence of a general factor, that Dr. Eysenck embarked on the investigation to which Mr. Paterson alludes. His final conclusion was that there is indeed "an objective general factor independent of teaching, tradition, and other irrelevant associations"³. Quite apart from Dr. Eysenck's research, there is a good deal of indirect evidence suggesting that the factor is, in part at least, innate; but for a more conclusive answer to Mr. Paterson's final question we must, I imagine, await the results of Miss Conway and others who are applying tests of various specialized abilities to twins who have been brought up separately almost from birth.

¹ Cf. P. Hartog, E. C. Rhodes, and C. Burt, *The Marks of Examiners*, 1936, esp. p. 306 and refs. The procedure described above had been used by Burt to eliminate the effects both of tested abilities and of hypothetical factors. But this device and the method of group factor analysis were both criticized by Prof. Pearson and Prof. Spearman and (for rather different reasons) by Dr. Rhodes and Dr. Stephenson.

² For the results see below pp. 91-92 (*Editor*). A convenient summary of the results obtained in this and later researches will be found in the section on 'Aesthetic Processes' in *Brit. J. Educ. Psychol.*, XIX, pp. 193f.

³ Mr. Paterson complains that Dr. Eysenck also fails to print a detailed table of the "correlations actually observed". But it should be remembered that, before the journal of statistical psychology was instituted, editors of psychological journals were reluctant to devote space and money to the printing of large and rather expensive tables of correlations, particularly as such tables were apt to evoke sharp protests from non-statistical readers who still looked on factor analysis as a somewhat dubious procedure and who were more likely to accept results based on the simpler methods described in elementary textbooks of statistics than those requiring large arrays of figures and elaborate methods of calculation which few in those days were likely to understand.

TABLE I. FACTOR SATURATIONS

(A). Williams, Winter, and Woods					(B). Burt and Pelling				
	G	L	V	M		G	L	V	M
A	·483	·414	—	—	L1	·556	·563	—	—
Bi	·632	·462	—	—	L2	·640	·488	—	—
Bii	·594	·508	—	—	L3	·523	·607	—	—
Ci	·518	·625	—	—					
Cii	·557	·391	—	—					
Fi	·545	—	·186	—	P1	·479	—	·531	—
Fii	·460	—	·263	—	P2	·594	—	·515	—
Fiii	·352	—	·230	—	P3	·450	—	·376	—
Fiv	·386	—	·098	—					
Gi	·369	—	—	·127	M1	·437	—	—	·382
Gii	·461	—	—	·081	M2	·281	—	—	·478
Giii	·293	—	—	·219	M3	·355	—	—	·431

EXPLANATION OF SYMBOLS

(A). Williams, Winter, and Woods :

1. Tests

- A. English composition
- B. Ranking extracts for literary merit : (i) prose ; (ii) verse
- C. Paired comparison : (i) prose ; (ii) verse
- F. Appreciation of pictures and photographs of household objects : (i) ranking pictures ; (ii) paired comparison of pictures ; (iii) ranking objects ; (iv) paired comparison of objects
- G. Appreciation of musical extracts : (i) ranking ; (ii) paired comparison ; (iii) triple comparison.

2. Factors

- G. general aesthetic factor ;
- L. special literary factor ;
- V. special factor for visual appreciation ;
- M. special musical factor.

(B). Burt and Pelling :

(See below p. 92).

THE GENERAL AESTHETIC FACTOR. III

By CYRIL BURT

At Miss Williams' suggestion I append the table of correlations and the factor saturations obtained in the experiments carried out by Miss Pelling and myself for the National Institute of Industrial Psychology. The details of the tests were not published at the time, since they were intended for private use in educational and vocational guidance. They were in fact very similar to those subsequently described and used by Williams, Dewar, and Wing in their later studies of literary, pictorial, and musical appreciation¹.

The subjects of the experiments were pupils attending or applying for admission to St. Martin's School of Art and other schools of art or evening classes under the London County Council, and applicants for vocational guidance tested at the National Institute.

¹ A brief account of the experiments and main results will be found in *Brit. J. Educ. Psychol.* XIX, pp. 193-94. For further details see references given there in footnote (1).

Tests of intelligence were also administered. In all 263 subjects were tested—the majority being between the ages of 17 and 25. None had I.Q.'s under 95 or over 125. The range of intelligence was therefore rather narrow ; and the elimination of differences in intelligence by partial correlation introduced comparatively slight changes in the correlations between the various aesthetic tests.

The correlations obtained after intelligence had been partialled out are shown in Table I and the factor saturations in Table II. Of the 78 correlations all are positive, and all except 5 statistically significant. The conclusions drawn were briefly summarized as follows : " Even after differences in general intelligence had been eliminated, a large factor still remained common to all the tests involving artistic appreciation or execution, . . . ' a general factor for aesthetic ability ' . This factor is not limited to ' art ' in the narrower sense, but enters into every manifestation of aesthetic taste—auditory or visual, verbal or concrete." Much the same view has since been expressed by Roger Fry : " If ", he writes, " we compare responses experienced in face of works of art of the most diverse kinds—architectural, pictorial, musical, or literary, we see in all these different expressions a general similarity in our mental attitude ; and further the attitude common to all these activities is also peculiar to them, and clearly distinguishable from our attitude in other experiences " (*Transformations*, 1926, pp. 1-2). And when both general intelligence and general appreciation have been eliminated, " a number of more specialized factors are found for the appreciation of (i) literature (prose as well as poetry) ; (ii) music ; (iii) painting, architecture, and sculpture ; (iv) ability to draw or paint, in short, for the aesthetic appreciation of different kinds of content¹."

Subsequent investigations furnished results which " seemed to imply the existence of specialized aptitudes, partly innate, for verbal, visual, auditory, or motor appreciation or skill. Cutting across these subdivisions according to content we also find two marked bipolar factors indicating a preference for (i) *classical* styles as contrasted with *romantic*, and (ii) *realistic* treatment as contrasted with *impressionistic*. These, however, appeared to be partly dependent on temperamental factors "².

TABLE I. CORRELATIONS BETWEEN TESTS

Test	L1	L2	L3	M1	M2	M3	A1	A2	P1	P2	P3	D1	D2
L1	(.694)	.688	.624	.333	.179	.454	.311	.121	.375	.302	.291	.155	.126
L2		.688	(.731)	.655	.424	.203	.396	.285	.134	.326	.267	.252	.263
L3			.624	.655	(.633)	.328	.396	.413	.209	.067	.247	.198	.203
M1				.333	.424	.328	(.582)	.518	.569	.247	.190	.341	.335
M2					.333	.424	.328	(.582)	.518	.569	.247	.190	.341
M3					.179	.203	.396	.518	(.584)	.552	.142	.156	.208
A1						.454	.396	.413	.569	.552	(.692)	.203	.134
A2							.311	.121	.375	.302	.291	.155	.126
P1								.311	.121	.375	.302	.291	.155
P2									.311	.121	.375	.302	.291
P3										.311	.121	.375	.302
D1											.311	.121	.375
D2												.311	.121

¹ There is in addition some slight evidence for factors underlying the different methods of testing that were used (ranking, paired comparison, and triple comparison) ; but the residuals suggesting these factors were too small to be statistically significant.

² The evidence for these further factors was given in *Character and Personality*, VII, pp. 291-299.

TABLE II. FACTOR SATURATIONS

Test	G	Factors			
		L	M	V	E
L1	·693	·441	—·063	·102	—·072
L2	·735	·409	—·078	—·083	·105
L3	·589	·520	·141	—·091	—·033
M1	·560	—·095	·468	·192	—·026
M2	·402	—·118	·634	—·075	·071
M3	·551	·213	·572	·094	—·082
A1	·328	·086	·036	·362	·097
A2	·233	·104	·027	·184	·054
P1	·395	—·137	—·043	·497	·271
P2	·451	—·128	—·014	·373	·128
P3	·320	·075	—·063	·415	·146
D1	·243	·073	—·057	·114	·512
D2	·279	—·073	·121	·068	·336

EXPLANATION OF SYMBOLS

1. Tests

L—Literary ability : (1) composition ; (2) ranked extracts ; (3) paired comparison of extracts.

M—Musical extracts : (1) ranked ; (2) paired comparison ; (3) triple comparison.

A—Specimens of sculpture : (1) ranked ; (2) paired comparison.

P—Reproduction of paintings : (1) ranked ; (2) paired comparison, (3) triple comparison.

D—Executive skill : (1) drawing ; (2) painting.

2. Factors

G—general aesthetic factor.

L—special literary factor,

M—special musical factor,

V—special factor for visual art,

E—special factor for executive skill in visual art.

STATISTICAL PSYCHOLOGY AND PHYSICAL LAWS

To the Editor, British Journal of Statistical Psychology.

SIR,—As no one better qualified has so far come forward to vindicate the claims of behaviourism, may I submit one or two obvious arguments in defence of those who, like Mr. Hansel, Mr. Willis, and Mr. Taylor, have advocated an essentially physical approach to the problems of psychology ? Mr. Taylor's chief contention, if I understand him rightly, is that the attempt to treat psychology as a statistical science has proved to be little more than 'a cloak for sterility'. Instead of examining the individual trees by the standard methods of physical science—observation and experiment, the statistical psychologist can see nothing but the forest. This curious omission is strikingly displayed in the section of Professor Burt's recent article, headed 'Planning Future Research' (this *Journal*, XII, esp. pp. 88–89), where only statistical methods are envisaged. It is equally obvious in his discussions of quantum theory in relation to psychology, where numerous equations are given (almost meaningless, I imagine, to most readers) and various novel concepts introduced which are left almost wholly unexplained. 'Causal hypotheses,' we are told, should be abandoned in favour of 'epistemological hypotheses'—based on such antiquated notions as 'consciousness', 'introspective contents', 'purpose', 'cognitive relations' and the like—vague metaphysical terms which have no place in natural science. The Newtonian or 'Galileian' approach, so ably championed by Lewin and Mr. Taylor

is discarded : the 'Aristotelian' is revived, with psychology, I suppose, to be defined once again as literally the science of the soul. May I therefore endorse in the strongest possible terms Mr. Taylor's dictum that " what psychology needs is not statistical laws but scientific laws ", i.e. generalizations based on the " universal laws of nature " ?

Mr. Hansel argues that the hypotheses of the psychologist must not contradict the ' laws of mechanics '. I venture to add a stronger requirement : the hypotheses of the psychologist must be *reducible* (at least in theory) to the ' laws of mechanics '. I use the word ' mechanics ' in the broad sense adopted by Professor Broad in his paper on ' Mechanical Explanation '.¹ The processes described and studied by the psychologist must (at least in theory) be reducible to those studied and described by the physiologist² ; these in turn to those of biochemistry ; and the chemical to the physical. The biologist has long ago got rid of the old metaphysical ideas of ' life ', ' vital force ', etc., regarded as influences or properties superadded to the physical ; and similarly the experimental behaviourist has got rid of the ideas of ' mind ', ' consciousness ' and ' psychic energy '.

There is thus a single all-embracing science based on a single set of postulates and concepts, namely those of physics. This conclusion indeed is so familiar³ that it is scarcely necessary to justify it at length. If Professor Burt and Mr. Gregory wish us to discard it, the onus is on them to prove it untenable : mere dogmatic dismissal is not enough.

The view I have set out here seems perfectly in keeping with the trend of modern psychology. Take, for instance, the most recently published textbook—Professor Bugelski's *Principles of Psychology*, which appeared this year. The definition of psychology and the whole treatment is essentially behaviouristic, and seems to me to follow much the same lines as Mr. Taylor has been advocating. Such terms as ' statistics ', ' probability ', ' consciousness ', etc., are not even mentioned in the index. The same standpoint is adopted by Mandler and Kessen in their discussion of *The Language of Psychology* (1959). " Physical terms," they say, " are the building stones of *all* scientific language." As they point out, this implies the total rejection of introspective (or ' phenomenological ') terms. Naturally the conclusions of both psychologist and physicist can be accepted only if they are confirmed by experience ; and by experience we mean the observation of scientists generally. That, of course, implies communication between one scientist and another. But to be communicated and to be testable they must be described in the *language* common to scientists. Statements referring to *private* experiences are thus ruled out.

Accepting then the familiar definition of psychology as " the scientific study of the behaviour, external and internal, of organisms " (cp. for instance, Bugelski, *loc. cit.* p. 20)

¹ *Aristot. Soc. Proc.*, XIX, pp. 86–124. Briefly his argument here is that " mechanical explanations imply that the phenomena in question obey Lagrange's equations " ; and this, he adds, does not mean that such explanations use no laws except the so-called " laws of motion ". The view I wish to defend is roughly that which Broad has very clearly formulated under the title of " mechanism " in his later book, *Mind and its Place in Nature* (pp. 44f). " This ", he says, " is a magnificent ideal, much more nearly true than anyone could have expected at first sight ; and the investigations pursued under its guidance have enabled us to discover many connexions which would otherwise have escaped our notice." The chief rival is " emergent materialism " to which Broad himself inclines, chiefly because secondary qualities (sound, colour, smell, etc.) are not reducible to physical. I should not object to this form of the theory, since it is essentially materialistic. ' Emergence ' merely means " not *at present* reducible " ; and many of the instances he cites from chemistry and biology have since proved reducible. As for secondary qualities, I should prefer the non-committal description ' epiphenomenal ', since they might prove to be merely ' parallel ' rather than ' emergent ' and we have no notion if or how they ' emerge '.

² Cp. C. C. Pratt, *The Logic of Psychology*, esp. chap. XIX.

³ Cp. the writings of contemporary behaviourists like Pratt, materialists like Hogben, physicalists like Carnap, and monists (neutral or otherwise) like Russell. If it be urged that the conception is an ideal or a methodological principle rather than a completed or accomplished fact, that would not affect my argument.

I suggest that ideally the only propositions that psychology should admit are of four main types :

(i) "At such and such a *place* (defined by the usual three spatial coordinates) and at such and such a *time* (t_0), such and such an organism, *O*, i.e. a specific type of physical system, is in such and such a *state* (again defined in physical terms)."

(ii) "At the *place* x_1, y_1, z_1 , and at *time* t_1 there is such and such a *stimulus*, *S* (source of light or sound, electrical stimulus to the cortex, etc.)."

(iii) "At *time* t_3 there is such and such an *external* (macroscopic) *response*, R_E (i.e. motions of specified parts of a specified organ, all defined by spatial positions)."

By way of 'intervening variables' we then hypothesize a physical explanation, viz.,

(iv) "At *time* t_2 there was such and such an *internal* (microscopic) *response*, R_I (a physical change in a physical part of the cortex)." If the subject, after being stimulated at t_1 , states that, between t_1 and t_3 (when he made the overt response), he "saw a bright light at x_2, y_2, z_2 ," we accept his verbal report but keep it within inverted commas, and interpret it as describing, not some mental event that happened in his personal and private consciousness, but (by a kind of code-language) some physical occurrence in the material structure of his brain, whether or not this was due to a direct electrical stimulus to that part of his brain or merely to an indirect stimulus from some external source of energy.

We thus obtain an observational statement that might be symbolized by a formula of the following type :

$$O(q_0) + S(q_1) \rightarrow R_I(q_2) \rightarrow R_E(q_3),$$

where the q 's will specify the four coordinates of space and time, O , S , and R are configurations of material particles (i.e. molecules) expressed, if more convenient, in terms of energies or forces, and the arrow indicates a deterministic relation of cause and effect. As a result of a number of such observational statements, we may hope to obtain one or more 'laws' which will have a similar form, but in which certain of the symbols will now denote generalized 'variables' instead of individual or particular 'constants'. This, I presume, is the type of 'law' that Mr. Taylor wishes the psychologist to seek instead of the ambiguous "statistical laws".

Now I take it both parties to the controversy are agreed that the first thing for a psychologist (including the statistical investigator) to do is to decide, and make clear, what is the general scientific background against which his investigations and his arguments are to be conducted. Mr. Taylor, Mr. Willis and Mr. Hansel, I take it, would all agree to my two main assumptions : (a) that there is now a satisfactory background consisting of a single comprehensive science and (b) that this science may be identified with 'Newtonian' science (to borrow Mr. Taylor's phrase) or, what amounts to the same thing, the 'laws of mechanics' (in Mr. Hansel's phrase) defined as above, or—to avoid verbal quibbles over the descriptive terms—let us say simply 'the laws of classical physics'. Professor Burt in his criticisms summarily dismisses the idea of such a unified scheme of science. My chief questions are these. Is his objection based on psychological grounds, viz., that such a science would have no place for any branch which dealt with mind or consciousness? Or is it based on non-psychological grounds, i.e. because the various branches of sciences do not in his view form a single compact and coherent system? If so, what are these irreducible branches of science? Are not his conclusions really the expression of a psychologist's opinions about physics rather than the conclusions of the vast majority of physicists themselves? Secondly, what are his reasons for proposing, as substitutes for the all pervasive laws of mechanics (or at least as supplementaries), such highly controversial concepts as 'fields', 'probabilities', the 'discontinuities' of the quantum theorist, the erratic phenomena said to be discovered in the world of subatomic physics, and—above all—statistical concepts and statistical laws in place of physical?

Consider first the concept of 'field'. A field is a purely abstract conception, like the lines of latitude and longitude. Science insists that all its causal concepts shall be of such a nature that they can be verified by direct observation, i.e. in terms of the actual

data of sense. Now we can no more perceive a 'field' than we can perceive the equator. On the other hand, matter is essentially perceptible, and therefore concrete. Once we have reduced some puzzling phenomenon to terms of matter—whether it be material atoms or molecules, or a material medium like the ether, or just the material tissues of the human body and brain—then we are explaining what happens by reference to the "basic substance of the universe". We are, in short, down at bedrock¹.

Or again why should we invoke such a controversial theory as that of 'general relativity'? Professor Burt and Mr. Gregory point out that this yields a better account of the perturbations shown by the orbit of Mercury, whereas the Newtonian theory of gravity, invoked by Mr. Taylor, leaves them unexplained. And we are apparently asked to believe that this is sufficient to dispose of the whole Newtonian theory of dynamics and the 'laws of mechanics' in term of which it is formulated. But is it not far more likely that some simpler explanation will be found to account for this trifling discrepancy?

Let us recollect that 'heat', and 'light', and 'electricity', which at one time formed separate and seemingly irreducible branches of physics, have each in its turn been reduced to micro-mechanical properties of material molecules (in the case of heat) or of a material ether (in the case of light and electricity). If, like electricity, life and mind have for so long remained a mystery, may we not now assume that, like electricity, they will soon be reduced to terms, not of some quasi-electrical field, but (as Sir Oliver Lodge has suggested) of "mechanical vibrations in the ether", or its modern equivalent?

After all, the speculations of Einstein and of Schrödinger can have no direct bearing on the investigations of the biologist or psychologist. For their purposes and ours all that is needed is to supplement the laws of physical science by (a) the fact that all living bodies evolve as a result of natural selection (a perfectly mechanical process) and so acquire adjustments, and (b) that certain living bodies, notably those with a nervous system, also acquire during their own individual lifetime still better adjustments as a result of 'conditioning' (also, it would seem, a mechanical process). This brings even 'psychological phenomena', i.e. the phenomena of human behaviour, within the sphere of a single all-embracing science. Does Professor Burt really wish to postulate anything further, apart perhaps from a possible preference for subatomic theories in place of the atomic theories of classical physics? And even were this further change accomplished should we not still remain within the realm of physics?

On these grounds therefore, I suggest, we seem bound to accept Mr. Taylor's contention that 'mind' is in fact simply "an integral part of an organic" (and therefore physical) "system, having the property" (shown by other physical systems, e.g. the atoms of chemistry) "of ultra stability". And, until these arguments have been refuted in detail, it is wholly irrational to expect us to reject behaviourism in favour of the obsolete concepts of introspective psychology, and so open up afresh "the wide gulf that formerly separated *Geisteswissenschaft* from *Naturwissenschaft*".

W. R. HAMMOND.

STATISTICAL PSYCHOLOGY AND PHYSICAL LAWS

REPLY

I fully agree with Mr. Hammond that "the first thing for a statistical psychologist to do is to make clear what is his scientific background". Statistical analysis is merely one accessory method for solving certain problems; and both the formulation of the

¹ I recognize that a 'field' with its 'lines of force' allows a neat and helpful diagram to be plotted of the forces which an actual test-object would encounter, if placed in a certain region of space, very much as the contour-lines on a map tell us what slopes or heights we might encounter if we climbed a particular mountain ridge. But the field-psychologist would apparently turn the imaginary contours into the reality and forget all about the solid hills and peaks.

problems and the interpretation of the results must always involve non-statistical considerations and must therefore be "fitted into the general context of current scientific opinion" as the investigator understands it¹. Nevertheless, I am not quite sure that a statistical journal is the appropriate place for a full discussion of the non-statistical aspects. Here, therefore, I shall try to answer Mr. Hammond's various questions as briefly as I can, and for details refer him to the relevant literature, either physical or psychological as the case may require².

First then do I dismiss the idea of "a single all-embracing science to be identified with physics"? And if so, do I dismiss it "on psychological grounds", i.e., because such a science "would leave no place for a branch which dealt with mind or consciousness"? If this means *solely* on psychological grounds, my answer is 'No'. What then are my "non-psychological grounds", and what are "the irreducible branches of science?" In replying to these further questions, to avoid Mr. Hammond's very reasonable suspicion that I am merely expressing "a psychologist's view about physical science", I shall, so far as possible, rely on the views expressed by contemporary physicists themselves—notably Bohr, Heisenberg, and Schrödinger³.

The notion that there is "a single all-embracing science, based on a single set of postulates and concepts, namely, those of physics" might have seemed almost a platitude in the days of Tait and Thomson's famous textbook. It still seemed plausible to those of us who studied elementary science fifty years ago. But it is certainly not true today. Instead of a single science I hold that there is a hierarchy of sciences, those dealing with higher orders of phenomena requiring laws and conception which are not reducible to the lower⁴. There is not just 'one closed system', but at least four systems or models.

(1) *Mechanics*. There is first the system of 'Newtonian mechanics' which Mr. Hammond alone accepts—the *système mécanique du monde* of Laplace, which treats the whole universe as one great self-operating machine, capable in principle of complete specification as regards past, present, and future independently of any conscious observer. For purposes of description it relies on the thing-language of everyday life, refined and reduced to the simplest possible terms, viz., (i) *particles* of matter having (ii) different positions in *space* at (iii) different instants of *time*. Time, space, and motion are continuous. All the individuals in the system (i.e., the particles) are exactly alike, and have only one intrinsic property, viz., mass. Stability is ensured by the law of the conservation of mass. Matter is thus devoid of quality, and hence both purposeless and senseless. In theory the 'true measurements' are exact; and the laws of motion certain. There is no need to invoke the notion of probability; and statistics, as Mr. Taylor rightly insists, form no part of the system. Since human beings themselves appear to be bodies made up of material particles, the first hypothesis to suggest itself to the lay psychologist is, very naturally, the assumption

¹ Cf. 'Definition and Scientific Method', this *Journal*, XI, p. 61. Mr. Hammond complains that, in my discussion of plans for research on telepathy, I mentioned statistical methods only. But I was there considering more particularly "the most satisfactory method of *demonstration*"; for other purposes I also emphasized the need for introspective analysis, laboratory research, and the study of spontaneous cases. I should indeed be the last to hold, as some investigators in the field of psychical research appear to do, that statistical analysis is the only type of worth-while investigation.

² Since in this country so many of those who have studied psychology have approached it through a training in Arts rather than in Science, I should perhaps apologize for the somewhat technical issues that I have touched on in this paper. To those who have already enjoyed a training in science, I must also ask forgiveness for what to them will doubtless seem a somewhat trite and superficial treatment.

³ N. Bohr, *The Atomic Theory and the Description of Nature* (1950), E. Schrödinger, *Science and the Human Temperament* (1958), W. Heisenberg, *Physics and Philosophy* (1960).

⁴ This in fact is the view taken by Broad (*Mind and its Place in Nature*, pp. 77).

that the individual person is merely a smaller machine within this vaster machine, much as Mr. Hammond and the earlier behaviourists have proposed. And in such a machine there would, I agree, be no room for any 'ghost'.

It is, I fancy, the fact that the whole system is expressed in the simplest possible terms of the ordinary thing-language that makes it so readily intelligible, and thus renders it attractive to the layman who has no time to delve into the subtleties of metaphysics or the later physical theories. For over two hundred years physics developed almost exclusively along the lines of Newtonian dynamics; and, in doing so, incorporated more or less successfully into this general framework the other branches of physics, which had hitherto been classified according to the human senses—namely, heat, optics, acoustics, and hydrodynamics, as well as mechanics in the narrower sense. Hence Hammond now wants it to absorb psychology.

(2) *Statistical Mechanics*. Difficulties, however, eventually appeared in the case of light and heat. To deal with the problems of heat a second closed system was gradually built up during the nineteenth century. In spite of the kinetic theory, it proved impossible to reduce thermodynamics to a special branch of mechanics. New concepts were required that had no place in the Newtonian mechanics, for example entropy; and in the mathematical treatment of the subject the concept of probability formed an essential element. Thus statistical methods in science already appeared indispensable.

At this stage too the traditional teaching of physics presented grave problems for the biologist and the psychologist. The second law of thermodynamics seemed to imply that the inevitable tendency of the physical universe (the passage from a less to a more probable state) entails a progressive breakdown of organized systems, with a dissipation of energy and an increase in general disorder; and it became difficult to understand how highly organized systems, like living bodies or conscious brains, could ever be built up. "So treibt uns der Mechanismus der Gegenwart dem Vitalismus der Zukunft mit Sicherheit entgegen."¹ And we find later writers, like Driesch, Bergson, and McDougall suggesting some supplementary concept, such as 'vital force' or 'psychic energy', to account for teleological organization². Yet, here again, I am in agreement with Mr. Hammond in rejecting any such concept—at least in the form in which those writers expounded it.

(3) *Field Theories*. A third system was developed to deal with the phenomena of electricity and magnetism, and rapidly extended to take in the phenomena of light and other forms of radiation, conceived at first as waves in the 'luminiferous ether'. The work of Faraday and Clerk Maxwell led to the fruitful notion of a 'field of force', the forces being directed quantities expressed by 'vectors' (a concept that meets us again in the factorial work of Pearson and Thurstone). The basic equations refer, not to the moving bodies, but to the fields themselves. In the later mathematical formulation the particles are reduced to singularities in the field; and, when the hypothesis of an all-pervasive ether was discarded, the fields were credited with the same sort of reality as was formerly been assigned to the material bodies or the particles³. At the same time the concept of force is replaced by the concept of energy—a somewhat elusive concept for the out-and-out materialist; energy becomes reified (the usual tendency of nouns in the thing-language);

¹ W. Bunge. *Vitalismus und Mechanismus*, 1886, p. 20. (Bunge, now almost forgotten, was one of the leading organic chemists of his day.)

² Cf. W. McDougall, *Modern Materialism and Emergent Evolution*, 1929.

³ The study of electro-magnetic phenomena gave the first clear indication of the inadequacies of a mere mechanical model. So far from providing an adequate explanation (as Mr. Hammond supposes), the mechanical theory issued in downright contradictions. Light and electro-magnetic waves are transverse, and transverse waves could occur in the ether only if the ether were solid: on the other hand, the ether offers no detectable resistance to the motion of the earth or other celestial bodies, and therefore cannot be solid. That these and other contradictions are inherent in the mechanical theory formed (as he himself tells us) one of the chief reasons that led Clerk Maxwell to undertake his celebrated investigations.

and physicists even talk of it as 'flowing' through space or spreading like a 'wave'. And now we find psychologists, like Ward, speaking of consciousness as a 'field', and others, like McDougall and Spearman, treating it as a form or manifestation of a special kind of 'energy'—possibly one of the forms into which physical energy may be transformed. Spearman indeed makes this conception the basis of a statistical theory, suggested it would seem by the statistical treatment of gases¹.

The paradoxes arising out of the notion of a stagnant ether were at length largely resolved by the theory of relativity. Einstein's 'special theory' introduced, not only a new conception of the 'field', but also new conceptions of space and time and the relation between the two, with corollaries which are quite incompatible with the Newtonian doctrine of an absolute time and an absolute space². Mass is no longer conserved, but can be transformed into energy. And thus slowly and reluctantly, the twentieth century physicist had to adapt himself to an entirely fresh conception of reality, quite unlike that envisaged by the physicists of the Victorian age or indeed by Mr. Hammond, Mr. Hansel, and Mr. Taylor.

(4) *Subatomic Physics*. Last of all, a new and yet completely coherent system is presented by the quantum theory, which also involves specific concepts and postulates, found in none of the other systems and in many respects incompatible with them. Here the probability function plays a still more important role. One far-reaching consequence is the changed conception of chemistry and still more of biochemistry—a change which Mr. Hammond entirely ignores. During the nineteenth century chemistry stuck obstinately outside the purview of physics. After Dalton had propounded the atomic theory numerous attempts were made to reduce chemical laws to terms of Newtonian mechanics applied to atoms, but failed completely, owing chiefly to the fact that the atoms of chemistry exhibit a stability which they could not possess on merely mechanical principles. As a result chemistry remained very largely a qualitative science. However, the stability of the atom can now be accounted for quite simply in terms of certain characteristics peculiar to quantum theory; and these characteristics, as it happens, are the very features which forbid a simple objective description of atomic structure in ordinary space and time. The quantum theory has thus reduced chemistry to a branch of quantum physics, though not of physics

¹ Mr. Hammond compares 'life and mind' to 'electricity', which for centuries remained a mystery, and suggests that "like electricity, they may soon be reduced, as Lodge suggested, to terms of mechanical vibrations in the ether" rather than to the properties of some "quasi-electrical field". But his interpretation of both Lodge's statements and mine is inaccurate. I myself had in view merely a tentative analogy which might yield a fruitful mathematical treatment; and the analogy was not with an *electrical* field but a *magnetic*. To put it simply but crudely, if 'life and mind' are to exert a guiding influence on a material world, then they must be able to deflect a material particle, such as an electron; and a magnetic field is of course constantly used to deflect streams of particles, such as the so-called cathode rays. But these fanciful speculations have little value except to show that the behaviourist's interpretation is not (as behaviourists apparently suppose) the only view compatible with modern physics.

² The theory of 'general relativity', though not so well established, nevertheless has far wider implications than the orbit of Mercury to which Mr. Hammond refers. It attempts to deal with two related anomalies which the Newtonian system leaves unexplained: (i) the fact that there are *physical* properties which enable us to distinguish between rotating and non-rotating systems of reference; (ii) the fact (which would be a pure coincidence according to Newtonian mechanics) that *gravitational* mass and *inertial* mass are always numerically equal. Einstein's solution is to deny, not only that space is absolute, but also that it is Euclidean. The theories of relativity (and the incidental corollaries, such as the phenomena of 'negative weight') plainly undermine all Mr. Hansel's *a priori* arguments for dismissing telepathy, precognition, and telekinesis as "incompatible with the teachings of physical science". Mr. Hansel is throughout manifestly postulating the Newtonian conception of space and time. Since no contemporary physicist today accepts that conception (except as a convenient working approximation) all his proofs fall to the ground.

as Mr. Hammond understands it. Nuclear researches have revealed forces far stronger—indeed, at times, we are told, over a million times stronger—than those previously encountered in mechanical physics. And within the nucleus the forces binding the particles are of a kind hitherto unknown : the complications involved in their study have raised problems still unsolved¹.

The nineteenth century notion of a single all-embracing science has thus broken down completely. Moreover, each of the systems I have mentioned is still far from final. At many points they seem inconsistent with each other. The assumptions of Newtonian mechanics are at variance with those of special relativity ; and the efforts made to combine the theory of special relativity with quantum theory have all hitherto led to mathematical impossibilities and contradictions (e.g., to a reversal of time or to negative values for the probabilities). In certain respects the first system can be regarded as a limiting case of the third, and to some extent of the fourth. This opens up the prospect that some day—as Pauli and Heisenberg have recently suggested—a fifth system might possibly be devised, of such a nature that each of the four existing systems could be deduced from it as limiting cases². But of one thing we may be assured. Such a system would certainly not consist in any theory or set of ‘ laws ’ that could be described in terms of our natural thing-language, or even of the concepts of classical physics : for they apply only to situations on approximately the same scale as those which are perceptible by our ordinary senses—situations in which the velocity of light can be treated as virtually infinite. Still less will it be a system of ‘ mechanical laws ’, even in the broad sense in which this phrase is used by Mr. Hammond and Mr. Hansel.

The Need for Supplementary Concepts. But, it will be said, what has all this to do with “ the problems of life and mind ” ? The speculations of Einstein and of Schrödinger about the exceedingly large or the exceedingly small, so Mr. Hammond believes, “ can have no bearing on the investigations of the biologist and the psychologist. For their purposes all that is needed ”, he tells us, “ is to supplement the laws of physical science by the facts (a) that all living bodies evolve as a result of natural selection (a perfectly mechanical process) and so acquire better adjustments, and (b) that certain living bodies, notably those with a nervous system, acquire during their own individual lifetime still better adjustments as a result of ‘ conditioning ’ (also, it would seem, a mechanical process).” In a word, the suggestion is that all we have to do is to add to what we know of the physics and chemistry of certain systems a knowledge of their history : biology and psychology can then be reduced to branches of the ‘ all-embracing science of physics ’. This view, however, was argued long ago in much the same terms by the nineteenth century forerunners of the behaviouristic school, notably by T. H. Huxley. The arguments then advanced were

¹ Mr. Hammond seems to forget that, in the transition from the mechanical theory to the quantum theory, the concept of a particle has itself been transformed. The so-called ‘ particle ’ is now something which possesses, not only mass, but charge, spin, and an enigmatic quality called ‘ strangeness ’, labelled like the others with a quantum number.

² It should be remembered that the three elementary bricks out of which subatomic theory was in its early stages constructed—the proton, the neutron, and the electron—by no means complete the list of elementary particles. More than thirty particles and anti-particles are known (not counting the most recent of all—the extremely hypothetical dualons or schizons). Of these some are conserved ; and some, strange to say, are not. What is still more remarkable is that all the elementary particles can, at sufficiently high energies, be transformed into others, or even created from, and annihilated into, energy. Matter therefore is no longer “ the basic substance of the universe ”. Nevertheless, we might plausibly conclude that each of the elementary particles is made up of the same “ substance ” (if we like to retain this questionable term), namely, energy : and in that case energy would become, as Heisenberg has pointed out, ‘ matter ’ in the Aristotelian sense—“ mere *potentia* which acquires ‘ actuality ’ by means of ‘ form ’ when the elementary particle is created ”. Heisenberg, it will be seen, does not share the low opinion which Mr. Taylor and Mr. Hammond evince for Aristotle and his followers, and would certainly reject Mr. Hammond’s conception of matter as the “ bedrock ”.

shown to be wholly inconclusive, and indeed self-contradictory, by Ward and other philosophic writers of the day¹. Mr. Hammond says nothing to rebut these objections, and in fact seems almost unaware of them. He goes on to ask whether I myself wish "to postulate anything further, apart from [my] preference for subatomic theories in place of the atomic theories of classical physics".

Here again, I think, it will be better to allow the physicist to suggest the answer. After all the physicist must know far better than either Mr. Hammond or I what are the limits (if any) of atomic or subatomic physics, and whether they are sufficient to account for the phenomena of "life and mind" without any further postulates. As regards the phenomena of life, Heisenberg, after citing specific examples, concludes that "there can be no doubt the laws of quantum theory play an important part in biological processes".² But he also contends that the degree of stability which living organisms manifest is far greater than anything which complicated structures consisting of many different types of molecules could be expected to display on the basis of physical or chemical laws alone: "Therefore something must be added to the laws of physics and chemistry before biological phenomena can be completely understood." It might, he believes, prove "necessary for any understanding of life to go beyond quantum theory, and construct a new coherent set of concepts, to which physics and chemistry could belong as limiting cases". Bohr has expressed much the same view: "our knowledge of the cell as alive may be 'complementary' to our knowledge, however complete, of the cell as a physico-chemical structure: moreover, since complete knowledge of that structure could be achieved only by operations that would destroy its structure, it may be a logical impossibility to make an exact determination of the structure of a living cell as such". The position in short, is not unlike that which arises in quantum physics, where an exact determination of both position and momentum simultaneously is impossible, and where the very attempt to fix the position of necessity changes that position³.

Both these writers state still more emphatically that an adequate account of psychological phenomena demands additional postulates. "If", says Heisenberg, "we go beyond biology and include psychology there can be hardly any doubt that the concepts of physics, chemistry, and evolution together will not be sufficient to describe the facts: on this point quantum theory has changed our attitude from that adopted by physicists of the nineteenth century." And, so far from rejecting mediaeval concepts, he goes on to suggest that the older concept of *Geist* (mind or soul) as found, for instance, in the philosophy of Thomas Aquinas is "far more natural and less forced" than the later doctrines which tend to hold that, if (as Descartes later supposed) animals are no more than physico-chemical machines, the same must be true of men.

Looking back at the different sets of scientific concepts formed in more recent times; we observe (he says) that they are in effect ordered by the increasing part played in the set by the subjective element. Classical physics was an idealization of the attempt to describe the world in complete separation from ourselves: relativity went far to undermine this assumption; quantum theory altogether forbids any completely objective description. And it is this increasing subjectivity which is in part responsible for the increasing use of statistics and the probability function.

¹ J. Ward, *Naturalism and Agnosticism* (1899).

² In the papers already cited I gave reasons for believing that both the processes determining inheritance and those occurring at the synaptic junctions are of the order studied by the quantum physicist.

³ Of course, were our problem, not so much to decide what hypotheses are theoretically permissible, but merely (as Mr. Hammond seems to hint in one of his footnotes) what working methods are to be preferred in the present stage of knowledge or ignorance, then everyone would readily agree that the proper line of research is to continue the attempt to explain as much as possible on the basis of accepted physico-chemical concepts before invoking anything new. But this is certainly not the problem as Mr. Hansel presents it.

Science, so I stated in one of the papers Mr. Hammond criticizes, no longer claims to describe 'external reality' as it is in and for itself (this *Journal*, XII, p. 126). Science is a human activity ; and can only describe the interplay between 'external reality' and the human observer. As both the theory of relativity and quantum theory have made clear, the observer and the observations are inseparable constituents of the 'reality' that science seeks to interpret. Classical physics assumed it was in theory possible to trace the course of any particle between two distinct observations. In quantum physics it is quite meaningless to ask what happens between one observation and the next. Nor can we predict with any certainty what will be observed at a later stage. It would, for example, be utterly impossible to trace in imagination the electron moving round in its orbit as we can trace the motion of the hour-hand of a watch ; it may jump from one orbit to another ; and, if anyone inquires where it is during the passage from one orbit to another, we cannot answer because such a question is devoid of meaning. In these respects the psychologist's observations of the workings of his own mind or of another's resemble the observations made in the quantum physics much more closely than they resemble those made in classical physics. Hence it is no argument against an alleged phenomenon to point out, as Mr. Hansel does, that its occurrence would contradict the 'laws of mechanics' : the behaviour of elementary particles themselves is constantly contradicting the 'laws of mechanics'.

Verification and Language. I am, however, in complete agreement with Mr. Hammond's argument that "the conclusions of both psychologist and physicist can be accepted only if they are confirmed by experience, i.e., by the observations of scientists generally", and that "this implies that, in order to be communicable and testable, they must be described in the language common to scientists". But I suspect my interpretation of these statements would be somewhat different from his. For instance, he seems to identify observations with "the data of sense".

Scientists are human beings ; and in their every day human life they are concerned with a world of things similar to that of all other human beings. That world is a perceptual world, not a world of sense-data. The discovery of sense-data is a late and comparatively rare achievement—the result of deliberate introspective analysis by a small and sophisticated group of scientists and philosophers—a type of analysis which Mr. Hammond strongly decries¹. The world of things, uncertain though it is in many ways, is primary ; and it is this anthropomorphic world that common language initially describes. Both our percepts and our modes of speech were evolved to enable us to meet our basic biological needs, i.e., to ensure survival. Nevertheless, when we rest from practical pursuits and contemplate this world of percepts, it proves to be far from self-explanatory or self-consistent. Man is therefore impelled to probe deeper in the hope of finding a truer world of reality behind this world of primary reality, which now seems to be merely a world of appearance. Down to the close of the nineteenth century this underlying causal world was still conceived as essentially a world of 'things' in the every day sense—a material world, a world of atomic billiard balls governed by the elementary laws of mechanics, capable in all its parts (according to Kelvin) of representation by a mechanical model, i.e., by a model made up of 'things' on the human scale. This, it will be seen, is still the implicit assumption on which most of the arguments of Mr. Hansel and Mr. Hammond are based.

Nevertheless, at some time or other the physicist's explanatory hypotheses with their manifold implications, however fantastic or farfetched, have necessarily to be related to this every day world of ordinary things. It would be impossible for the quantum theorist to carry out his experiments unless the particular system he was studying interacted with the larger system of which he as observer forms a part. Moreover (and here, I fancy, is a point which is all too frequently overlooked), *this wider system is always described even by*

¹ The trained eye of a painter or an observant psychologist is required to see the true colours of the visual sense-data that make up a landscape : the familiar experiment of looking at a distant view with the head upside down is alone sufficient to reveal the difference between what is sensed and what is perceived.

the quantum physicist in the ordinary thing-language, or at any rate in the thing-language of the classical physicist. The pointer and its readings, the screen and its scintillations, the Geiger counter and its clicks, the tracks in the cloud-chamber, the photographs obtained with the spectroscope or the electron-microscope—these are referred to or reported, not in terms of imperceptible particles or mathematical symbols, but (as the stock descriptions I have just used sufficiently show) in terms of perceptible things.

And the same will be true of the psychologist. All his theories, even when they relate to the highly abstract or hypothetical concepts of the neurologist, the psychoanalyst, the spiritualist, or the statistical computor, must in the end be related to, and confirmed by, ordinary human behaviour or ordinary human consciousness. This is a requirement which the statistical psychologist is very apt to forget. By a 'logical construction' he builds up his 'factors' or his hypothetical 'abilities' out of his initial data; and, more often than not, when he comes to interpret or name them, he omits the final observations needed to link them up with actual behaviour or actual mental processes. Moreover, since (except for our own introspection) what we know of conscious experience can only reach us in the form of verbal reports from our subjects or from other psychologists, these too will be in one sense items of behaviour. It is, I suspect, these imperative requirements that have lent plausibility to behaviouristic doctrines. But to infer from this, as the behaviourist does, that psychology is *merely* "the study of behaviour internal or external" is like inferring that quantum theory is concerned *merely* with the phenomena of classical physics, because the language of classical physics is always used to describe the conditions of the experiment and the results as actually observed.

Science is based on Successive Approximation. The great mistake of Mr. Hansel and Mr. Hammond, as it seems to me, arises from their belief that the 'laws of mechanics' are exact, final, and sacrosanct—at any rate so far as their own particular area is concerned; whereas they represent no more than a first approximation. Our interpretation of the world develops by series of stages, and is always subject to progressive revision. This is manifestly true of the four scientific systems we have been discussing: each in some measure rectifies or reduces the errors of its predecessor, and extends its scope, at the same time becoming more and more comprehensive and more and more complex. We have analysed the material things of everyday life into their basic constituents; we have reduced these constituents to their component molecules; we have penetrated the molecule and found the atom; we have penetrated the atom and found the elementary particle—each new advance bringing with it a theory of ever-increasing complexity. The next step presumably will be to squeeze inside the elementary particles¹. And, as we step down from the realm of Lilliput and enter the Alice-in-Wonderland universe that apparently awaits us at this level, we must be prepared to encounter events and possibilities (we can no longer call them 'things' or 'laws') still more remote from the world of classical physics than the subatomic world already revealed by relativity and quantum theory.

A similar progress from the simple to the complex is evinced as we move from the elementary science of mechanics and dynamics to chemistry, biology, and finally to psychology—each new advance bringing our knowledge nearer to the true richness of the universe, and each requiring new categories or concepts. Let us note, too, that the change

¹ Here Heisenberg and others have made an instructive suggestion. To determine the scale of the particular phenomena which a physicist selects for his investigations, at least three fundamental units are required—e.g., the centimetre, the second, and the gram if we decide to choose length, time, and mass. We could, however, just as well take length, velocity, and energy. Now in subatomic physics the velocity of light and Planck's constant provide two possible units. For the third 'dimension' an appropriate choice would be a new unit of length of the order of about 10^{-13} (the radius of the lighter atomic nuclei). If from these we deduce an expression corresponding to a mass, the magnitude turns out to equal to that of an elementary particle. But, when we attempt to follow the mathematical consequences of this and the other modifications necessitated by this change of level, all kinds of paradoxical corollaries apparently ensue. Curiously enough, much the same paradoxes appear when we analyse psychological measurements.

is not so much a change in the type of entities investigated (as the popular definitions suggest), but rather in the types of relation and structure to be included in our study. If therefore the psychologist desires to bring in new relations such as cognition (i.e., simple awareness) and possibly pre-cognition and extra-sensory cognition, or it may be teleological relations as distinct from causal, it would be a fatal error to ban them just because they are not reducible to the mechanical relations studied at the simpler stage.

The Need for Statistical Techniques. It is at these more complex levels that statistical methods become essential. They are required partly because we are dealing, not with isolated objects but with large populations, partly because there are a large number of variables varying at once, and partly because, when we transcend the strictly causal determinations of the simpler and more abstract systems, the element of probability plays an even larger role. All this Mr. Hammond's formulae entirely ignore. I confess that I am very doubtful whether it is possible in advance to dictate the kind of mould into which all permissible propositions in a given science can be cast, or even to fit the propositions already validly established into one general scheme. However, if Mr. Hammond insists, then in place of his causal formulae, I should prefer one of the following basic type :

$$(x, y) (\Phi x \overset{p}{\supset} \Psi y),$$

which might be translated (not quite accurately) : "for every corresponding pair of x and y , ' x is Ψ ' implies, with a degree of probability p , that ' y is Φ '." This formulation differs from his in three main ways. (i) Instead of categorical propositions asserting that ' y is certainly Ψ ' or the like, I hold that, in an empirical science, only *probable* statements are allowable : (this of course rules out all assertions about causality in the deterministic sense). (ii) Instead of limiting the basic formulae to subject-and-predicate statements, my notation is intended to include relational propositions. (iii) If the statement is to form an essential part of the science of psychology, then, instead of limiting the meaning of the symbols to physical systems and physical properties, one or more of them must refer (directly or indirectly) to some conscious system, or some conscious quality, or some conscious relation¹.

CYRIL BURT.

¹ In defence of the standpoint I have here very briefly and inadequately outlined, I must refer to my previous papers (e.g., 'Definition and Scientific Method in Psychology', this *Journal*, XI pp. 31-69). As regards notation, Φ and Ψ are what some logicians call 'functors', i.e., signs denoting functions—e.g., relations of various types as well as predicates in the narrower sense : (this is why the verbal translation given in the text is not quite accurate). As illustration of the kind of propositions I have in mind the following may perhaps serve. "If the back of my hand is stimulated with a hair-aesthesiometer at a pressure of 1.6 gr/mm, there is a probability of 0.5 that I shall feel a sensation of touch at that place." "If a male adult picked at random talks to me with a sound-intensity averaging 25 decibels at the ear, there is a probability of 0.25 that I shall correctly understand what he has said." "If x is a boy of school age with an I.Q. of 80, and y his brother, there is a probability of 2 to 1 that y will have an I.Q. between 78 and 96."

A NOTE ON THE INTELLIGENCE OF TWINS

By J. CONWAY.

Maddox and others have recently suggested that the low average intelligence of twins may be due to the fact that the prenatal conditions are often unfavourable (cf. this *Journal*, XII, 1959, p. 135). And they go on to argue that, since the varying intra-uterine conditions must be of the same severity for both members of a pair, that would suffice to explain the higher correlations for intelligence reported in my previous articles.

But plainly the type of causation assumed would affect dizygotic twins in much the same way as monozygotic. Yet the correlation for the former is not appreciably higher than the correlation for ordinary siblings. The mere fact that monozygotic twins are enclosed in the same membrane, as well as in the same uterus, could not possibly account for the much higher correlation obtained in their case. Moreover, if the correlations are due solely to similarity of uterine environment, how does Mr. Maddox account for the correlation between the twins and their parents?

However, the assumption that the uterine conditions are the same for both members of a pair is of course incorrect. For one thing they cannot both occupy the same position. To what extent the *effects* are the same for both is a question for investigation. In the L.C.C. institutions for mental defectives, where twins were studied by Professor Burt and myself, the *physical* effects were not the same: usually the second to be delivered suffered most. Further, out of 26 cases where low grade mental deficiency occurred in a monozygotic twin, in 24 the other twin was *not* mentally deficient: (this agrees with the findings reported from other institutions). With high grade or borderline defectives, unaccompanied by physical defect, the opposite is the case. Outside the institutions the amount of mental inferiority showed only a low correlation with the amount of physical damage. It would seem therefore that, when the usual intra-uterine conditions of twins do have any effect, they are more likely to reduce the correlations than to increase them.

BOOK REVIEWS

Educational Guidance and the Pool of Ability. By D. M. McINTOSH. University of London Press, 1959. Pp. 200. 21s.

Of the many investigations sponsored by the International Institute Examinations Inquiry few have aroused a wider interest, particularly among Scottish educationists, than the research carried out by Professor McClelland and his colleagues on *Selection for Secondary Education*. Broadly speaking it might be said that McClelland's investigation at Dundee did for the Scottish system very much what the L.C.C.'s psychologist in London, and Birmingham, and elsewhere did for the English system. Both sets of inquiries were preceded by surveys of the general school population in the area; and both were the outcome of a conviction that the methods of allocating children to secondary schools at the age of eleven missed far more pupils of high ability than they actually selected. Both, too, proposed much the same scheme for detecting "the most deserving candidates", namely, a composite battery of so-called 'intelligence' tests, a standardized test of attainments in English and in Arithmetic, and a series of personal assessments submitted by the children's teachers, rescaled by the aid of the test to allow for variations in standard from one school to another¹.

Professor McClelland's chief assistant was Dr. McIntosh. He was subsequently appointed Director of Education for Fife, and has now published a research which he himself has inaugurated in his own education area. In the Dundee inquiry the training of the testers, and the marking and statistical calculations, were carried out largely by the staff and students of the Training College. However, educational research organized by members of an academic department necessarily suffers from serious restrictions which are not imposed on an inside investigator who, like Dr. McIntosh or the L.C.C. psychologist, is a member of the education authority's own staff. In this and other ways Dr. McIntosh's investigation thus marks a definite advance on the pioneer inquiry conducted by Professor McClelland. And, although the professional statistician and psychologist may be tempted to criticize minor details here and there, his general approach to the problem will probably command the attention of the practical teacher and education officer in a way which the recent reports of the British Psychological Society and of outside workers have apparently failed to do. His whole discussion is imbued with a first-hand understanding of the practical problems involved, such as can only be gained by actual experience within an education department.

In the earlier inquiry at Dundee an undertaking had to be given that the experimental data would play no part in the education authority's allocations. In the new inquiry the data used are those of an actual promotion scheme; and it has been possible to obtain the educational history of each child throughout the whole of his secondary school career. The Dundee investigation was restricted to urban schools; in the county of Fife almost every type of school is represented. In 1935 pupils at secondary schools paid fees; today education is free. Finally, in many important ways the educational, social, and economic conditions have drastically changed.

Dr. McIntosh's primary aim was "to test the efficiency with which pupils are allocated to appropriate secondary courses under the scheme of allocation recommended by the Advisory Council on Education in Scotland".

The pupils for whom criterion-estimates were available necessarily formed a selected group. Hence, in assessing the predictive efficiency of the initial assessments it was requisite to make some sort of correction for the effects of selection. The formulae

¹ Burt also recommended *two* tests of intelligence applied at different dates, and for borderline cases a consideration of home circumstances and character qualities in addition to supplementary tests at an interview: cf. 'The Selection of Pupils for Different Types of Secondary Schools', *Brit. J. Educ. Psychol.*, XVII, pp. 57-71, XX, pp. 1-10.

commonly used are those of Karl Pearson. Dr. McIntosh apparently has adopted much the same procedure with minor modifications. Thus corrected the correlation between the results of the initial assessment at the time of allocation and the final result of the Scottish Leaving Certificate Examination is stated to be 0.87—a distinctly high figure. The correlation for the teachers' assessment for English was almost as high. The poorest results were those furnished by the test and estimates for arithmetic.

A supplementary investigation was undertaken into the problem of wastage. This was found to be nearly twice as high among girls as among boys. The chief causes are said to have been physical defects, adverse home conditions, lack of interest in the type of education provided, the competing attractions of lucrative employment and leisure, and above all defective qualities of character. Measures are suggested which might do much towards mitigating those unfortunate tendencies.

Perhaps the most instructive parts of the book are those concerned with "the actual supply of intellectual man-power" and "the methods of making the most of whatever is available". Dr. McIntosh accepts without hesitation the doctrine that the capacities of different individuals are dependent on their genetic constitution. "The conception of a limited pool of ability," he says, "is at last obtaining tardy recognition: success in the struggles which are now taking place can be achieved only by those nations which are prepared to develop and deploy the full potential of their intellectual resources." As a convenient definition for those who are to be included in the 'pool' he suggests the attainment of at least two higher and three lower passes in the Leaving Certificate examination (or its equivalent). And on this basis he finds that just over 11 per cent. of each group would be available to form the pool he has in mind. Dr. McClelland's estimate was rather more generous—about 16 per cent. In England nearly twice the proportion—well over 20 per cent.—are now receiving secondary education of a grammar school type. But two points should be noted. First, as Dr. McIntosh observes, the home conditions and the parents' attitude probably prevent many children from developing their intellectual potentialities to the full; moreover, the secondary courses still have a strong academic and bookish bias, whereas the abilities of many intelligent pupils lie in a more practical or technical direction. Secondly, owing to the inevitable imperfections of even the most efficient methods of selection, the only way to make certain of securing the highest 10 or 12 per cent. is to select at least twice that number.

Dr. McIntosh's report of the whole investigation thus provides an illuminating and up-to-date contribution to a most urgent national problem. It is full of incidental points of special interest to teachers and educationists, and well deserves the attention both of psychologists and of those engaged in educational administration. J. W. BROOK.

Rapid Statistical Calculations. By H. M. QUENOUILLE. London: Charles Griffin & Co., 1959. Pp. xv+96. 10s.

In this useful pocket manual Mr. Quenouille, a member of the Research Techniques Unit organized by Professor Kendall at the London School of Economics, has collected a set of "distribution-free and easy methods of estimation and testing". As he points out, a quick method is no longer quick when the computer has to spend some time hunting it up and finding out whether or not it is really appropriate to the problem in hand. Accordingly he has included in his selection only those procedures which in his view are of the greatest and most frequent use in practical work; and has arranged his exposition so that two facing pages are devoted to each method, the explanation on the left-hand page and the worked example on the right-hand page. For each method he states in succinct terms what precisely it does, how it does it, what degree of accuracy is to be expected and how much is lost, and what assumptions are implied in adopting the method. Of the thirty-five methods described about half are concerned with the calculation of mean level

and scatter for groups of observations, four with the comparison of frequencies and proportions, and the remainder with the calculation of association. The appendix gives tables of the usual types—square roots, normal deviates, significance levels, and the like. The book concludes with a valuable list of references, partly general and partly giving fuller details for particular procedures.

Both workers and teachers in the field of statistical psychology will find this handy little volume invaluable, and to many of them several of the methods will probably be entirely new.

CHARLOTTE BANKS.

Experimental Design in Psychological Research. Revised Edition. By ALLEN L. EDWARDS. New York : Rinehart & Co., 1960. Pp. xvi + 398. \$ 6.50.

Professor Edwards' well-known textbook has been thoroughly revised and many of the chapters completely rewritten. A new chapter on 'Trend Analysis' has also been included. For the rest the principles followed are much the same as in the first edition. The book is intended to be intelligible to those who have a knowledge of elementary algebra and the general rudiments of statistical analysis. As before, each chapter ends with a series of 'examples' for the student to work out ; but with these are now included a number of questions intended to test the pupil's knowledge of concepts and theoretical principles. The page is slightly bigger ; and the type slightly smaller. As a result, although the book itself is somewhat enlarged, the number of pages is fewer ; and the whole format is much improved.

In the main the experiments in whose design the psychologist is assumed to be chiefly interested are those requiring analysis of variance or some closely related procedure. Factor analysis is nowhere mentioned. The recent paper by Burt and Gregory on 'Scientific Method in Psychology', which appeared in this *Journal*, is cited ; but otherwise work by British statistical psychologists is wholly ignored. Nevertheless, within the field to which the author restricts himself, the volume unquestionably remains one of the clearest and most helpful textbooks.

F. N. HARPER.

Turning Points in Physics. Edited by A. C. CROMBIE. Amsterdam : North-Holland Publishing Co., 1960. Pp. 192. 12s.

At Oxford, during recent years, it has become the practice to arrange, as part of the courses in the history and philosophy of science, a series of lectures which shall present a critical discussion of important aspects of contemporary scientific thought. The present series was delivered during the Summer Term of 1958.

Dr. Crombie has contributed a brief but admirable introduction summarizing the general reasons for the revolutionary changes that have taken place in physical theories during the past half century. The first chapter, by Dr. Blin-Stoyle, describes "the end of the mechanistic theory and the rise of field physics". Dr. Tel Haar then explains "the quantum nature of matter and radiation" ; and this is followed by two important and illuminating papers—by Dr. Mendelsohn on the "entrance of probability into the physical sciences", and by Dr. Waismann on "the decline and fall of causality" (the longest in the book). Professor Temple discusses from a somewhat novel standpoint "the theory of relativity" ; and Professor Wilkinson ends the series with an analysis of the "new concepts".

Except perhaps for the final chapter, which is somewhat technical, the book should be intelligible even to those who, like many British psychologists, are comparatively innocent of the basic sciences which have nowadays begun to influence their own. It may be specially commended to those who (as several contributors to this *Journal* show) are specially interested in the bearing of scientific method and of physical concepts on what

they call the 'behavioural sciences'. Here, instead of constant references to continental workers who write in a foreign language, the latest views on physics are expounded by British scientists. Every contributor in fact is a member of the University of Oxford. The problems raised are of profound psychological as well as philosophical importance. The book itself is delightfully printed and excellently produced. W. L. GREGORY.

Readings in General Psychology. Edited by P. Halmos and A. Aliffe. London : Routledge & Kegan Paul, 1959, Pp. x+251. 25s.

The majority of textbooks written for the student of elementary psychology attempt to present a comprehensive survey of the whole, and as a rule have necessarily to adopt a somewhat superficial type of treatment. The aim of the editors of the present volume has been to compile a different type of introduction which shall offer a more intensive examination of a limited number of all-important topics.

After an introductory chapter on the 'scientific status' by one of the editors, the series begins with two physiological papers, on cortical localization by Adrian and electroencephalography by Osselton. Burt's work is represented by his article on the 'meaning and assessment of intelligence', and Vernon's by one on 'interpreting test-results'. Inheritance and extraversion are discussed by Eysenck, instinct by F. V. Smith, and personality by Bowlby. These are followed by an examination of the concepts of the normal mind, of role and status, and of group dynamics. Oldfield discusses skill, and Hearnshaw the psychology of thinking. The whole concludes with a paper from the other editor on 'psychology and ethics'. Most of the contributions have already appeared in various journals. Although one (by Ernest Jones) goes back to 1931, the majority are recent and up to date; all are by British writers, and are of at least as high a standard of excellence as the American extracts which have become so fashionable—and far more readable.

E. F. JACKSON.